Essays in Industrial Organization

by

Chung-Ying Lee

Department of Economics
Duke University

Date: ______________________

Approved:

____
Andrew Sweeting, Supervisor

____
Peter Arcidiacono

____
James W. Roberts

____
Wilfred Amaldoss

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2014

Abstract

Essays in Industrial Organization

by

Chung-Ying Lee

Department of Economics
Duke University

Date: ______________________
Approved:

____
Andrew Sweeting, Supervisor

____
Peter Arcidiacono

____
James W. Roberts

____
Wilfred Amaldoss

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Economics in the Graduate School of Duke University
2014

Abstract

The dissertation consists of three chapters relating to pricing strategies. Chapter 1 studies coupons for prescription drugs and their impacts on consumer welfare, firm profits, and insurance payments. Chapter 2 examines consumer brand loyalty and learning in pharmaceutical demand and discusses implications for marketing and health care policy. Chapter 3 develops a framework for estimating demand and supply in an online market with many competing sellers and frequent price changes and proposes optimal pricing strategies for a large seller.

The first chapter studies an innovative price strategy in pharmaceuticals. Branded drug manufacturers have recently started to issue copay coupons as part of their strategy to compete with generics when their branded drugs are coming off patent. To explore the welfare implications of copay coupons, I estimate a model of demand and supply using pharmaceutical data on sales, prices, advertising, and copayments for cholesterol-lowering drugs and perform a counterfactual analysis where a branded manufacturer introduces coupons. The model allows flexible substitution patterns and consumer heterogeneity in price sensitivities and preferences for branded drugs. The counterfactuals quantify the effects of copay coupons for different assumptions about the take-up of coupons and the ability of the branded manufacturer to direct them to the most price-sensitive types of consumers. The results show that the agency problem between insurers and patients gives a branded manufacturer a strong incentive to issue copay coupons. Introducing copay coupons benefits the

coupon issuer and consumers but raises insurance payments. In equilibrium, insurer spending can increase by as much as 25% even when just 5% of consumers have a coupon, with social welfare falling significantly.

Medicines for chronic conditions like high cholesterol, heart disease, and diabetes are repeatedly used for a long period of time. Consumer dynamics thus plays an important role in the demand for those drugs. In the second chapter, I estimate a demand model with brand loyalty and learning using micro-level data from cholesterol lowering drug markets in the United States. The estimates suggest high switching costs and strong learning effects at the molecule level in the markets. Switching costs raise the predicted probability of choosing the same drugs in a row and learning largely increases patient stickiness to a molecule in the long run. I discuss pricing implications of the estimation results for drug manufacturers, insurance companies, and policy makers.

The last chapter, coauthored with Dr. Andrew Sweeting and Dr. James W. Roberts, looks at pricing in a different context. We estimate a model of entry, exit and pricing decisions in an online market for event tickets where there are many competing sellers and prices change frequently. We use the estimates from our model to analyze the optimality of the pricing policy used by the largest seller (broker) in the market. We show that the broker's pricing policies substantially affect the prices set by his competitors. When we compare the broker's pricing policy with the prices that our model predicts are optimal we find that the broker sets approximately correct prices close to the game, when his pricing problem resembles a static one, but that he might be able to gain from using different pricing rules and updating prices more frequently further from the game.

To my beloved parents.

Contents

List of Tables

List of Figures

Acknowledgements

I would like to deeply thank my adviser, Andrew Sweeting, as well as my committee members Peter Arcidiacono, Jimmy Roberts, and Wilfred Amaldoss for their invaluable guidance and support throughout my work on this project. I am also grateful to all seminar and lunch group attendants at Duke University for their insightful comments and helpful suggestions.

1

A Welfare Analysis of Copay Coupons in Pharmaceuticals

1.1 Introduction

Coupons have long been prevalent in consumer goods, but recently they also started to play an important role in the pharmaceutical industry. The coupons distributed by drug manufacturers, called copay coupons or copay cards, reduce consumers' out-of-pocket costs of prescription drugs. Many top-selling drugs, including cholesterol fighter Lipitor, blood thinner Plavix, and blood pressure drug Diovan, started to offer copay coupons as they were coming off patent in recent years. Analysts estimate about 13% of branded prescriptions were associated with copay coupons in 2011.[1] The number of prescriptions filled with copay coupons is expected to increase approximately 15% per year.[2] Despite the fast growing use of copay coupons, little empirical work has been done to examine the welfare implications of this new strategy. In this paper, I provide a counterfactual analysis of how copay coupons affect

[1] "Copay Cards Could Be Win-Win if All Sides Work Together," *Drug Benefit News*, July 22, 2011.

[2] "How Copay Coupons Could Raise Prescription Drug Costs by $32 Billion Over the Next Decade," Pharmaceutical Care Management Association, November 2011.

consumer welfare, firm profits and insurance payments, using a model estimated with data from the market for cholesterol-lowering drugs.

The study focuses on a new incentive to issue coupons: the agency problem between insurers and patients in pharmaceuticals. Coupons are widely used by firms in other industries to compete for price-sensitive consumers. By distributing coupons in a market, firms can rely on consumer self-selection to achieve market segmentation and increase profits. Narasimhan (1984) shows that coupon users are more price sensitive than nonusers and firms could thus offer coupon users a lower price and raise the price for nonusers, allowing for price discrimination. The special market structure in the pharmaceutical industry gives drug manufacturers another incentive to use coupons. For prescription drugs, doctors and patients make the purchase decision and insurance companies pay for most of the drug costs. Insurance companies in the U.S. usually ask for a lower cost share from patients for less expensive drugs to reduce spending. By issuing copay coupons directly to doctors and patients, drug manufacturers can lower the out-of-pocket cost for patients and induce them to choose the drugs with coupons even though this may increase the cost to the insurance company. In this paper, I use counterfactuals to explore both the effects of this agency issue and how the ability to target coupons to particular types of consumer can affect the profitability of this strategy.

The welfare effects of copay coupons hinge on the substitution patterns in pharmaceuticals. To capture the key features in the markets, I estimate a model with rich substitution patterns and consumer heterogeneity in both price sensitivities and preference for branded drugs, using unique datasets on sales, advertising, and co-payments. The model captures competition among drugs in various ways, including classes, molecules, forms and brandedness. Also, the model allows consumers' price sensitivities and preference for branded drugs to be drawn from a binary distribution. The consumer heterogeneity helps to explain why branded drug prices usually stay high after patent expiration. The estimates show that (1) the substitution is strongest among drugs with the same molecules, and (2) consumers who have a preference for branded drugs are less price sensitive.

The counterfactual results show that copay coupons have a large impact on social welfare even when only a small fraction of consumers get them. I simulate the outcomes of a copay coupon program introduced by the manufacturer of a branded cholesterol-lowering drug after patent expiration. I consider different assumptions about the take-up of coupons and the ability of the branded manufacturer to direct them to the most price-sensitive types of consumers. In the baseline case, coupon users and nonusers are equally price sensitive. In the targeting case, coupon users are more price sensitive. I find that, in the baseline case, the firm that introduces copay coupons would set a very low copay for coupon users and raise the full price to reap a large profit from insurers. In equilibrium, consumers gain from the lower copay by using coupons but most of the other drug manufacturers' profits shrink as the coupon program expands. Insurance payments increase by 25% when only 5% of consumers have a coupon. The net effect of copay coupons on social welfare is negative due to the large increase in insurance payments.

In the targeting case, copay coupons help to segment markets when the branded manufacturer offer coupons to more price-sensitive consumers. Most of branded drug manufacturers make a larger profit by exploiting the lower price-sensitivity of those without a coupon and charging them a higher price. The copay coupon thus mitigates price competition among branded drugs and improve their profits. At the same time, the higher prices faced by those without a coupon lowers the overall consumer welfare gain. The net effects of copay coupons on social welfare do not change much when I include the ability to target coupons to price-sensitive consumers.

The paper contributes to the literature on couponing by considering a new incentive for using coupons. Theoretical work by Holmes (1989) and Corts (1998) show that offering coupons could help price discrimination in an oligopoly setting since coupons attract price-sensitive consumers. Under certain conditions, such price discrimination will lead to higher prices and profits. Narasimhan (1984) empirically finds that coupon users are more price-sensitive than nonusers. Another empirical study by Nevo and Wolfram (2002) shows that coupons can spur price competition and lower shelf prices when they are widely available. Unlike typical coupons, copay coupons in pharmaceuticals can lead to higher drug prices because of an agency problem between insurers and patients. In this industry, consumers or patients share only a small portion of drug costs, and I show how this tends to increase the profitability of a coupon.

In addition, the structural model in the paper incorporates two important features for pharmaceutical demand estimation. First, I apply a generalized extreme value (GEV) model developed by Bresnahan et al. (1997) to capture substitution patterns along different dimensions. Arcidiacono et al. (2013) use a similar model to study the welfare impacts of me-too and generic drugs. The model has a nesting structure that allows for consumer switching based on different drug characteristics, including molecule, class, branded/generic, and form. This strength facilitates simulation of introducing copay coupons since the model captures how consumers' choices change given the new pricing strategy. Second, following Berry et al. (2006) in their analysis of airline industry, I consider two types of consumers who differ in their price sensitivities and choice sets. The consumer heterogeneity could help to explain branded drugs' pricing after they lose patent protection. Also, the consumer heterogeneity makes it possible to separate the motivations of couponing. In the two-type framework, I can control for coupon users' type and examine how the agency problem between insurers and patients would affect equilibrium prices when copay coupons are used.

Finally, the unique datasets used in the demand estimation of this paper cover major aspects in pharmaceuticals, including prescription drug sales, physician advertising, direct-to-consumer advertising (DTCA), and copayments. Jayawardhana (2013) is one of the few empirical studies that use these rich data in demand estimation for pharmaceuticals. Most of the other papers in pharmaceutical literature use a single source of advertising to represent the marketing from drug manufacturers and/or include full prices in demand. In the case of cholesterol lowering drugs considered in this paper, DTCA and physician advertising both play an important role in marketing and adding them can help to explain substitutions among drugs. Additionally, copayments are the actual prices faced by consumers. Using full prices in demand estimation would underestimate the price coefficient. The data on copayments also helps to build the relationship between full prices and copayments and facilitate counterfactuals in which firms change full prices to affect insurance copayments and demand for drugs when copay coupons are introduced.

The rest of the paper is organized as follows. Section 1.2 provides industry background and relevant information about copay coupons. Section 1.3 describes data. Section 1.4 develops models for demand, supply, and copayments. Section 1.5 discusses estimation strategies and estimation results. Section 1.6 presents counterfactuals for introducing copay coupons under two scenarios. Section 1.7 concludes.

1.2 Copay Coupons in Pharmaceuticals

Copay coupons are instantaneous rebates to patients usually offered by branded drug manufacturers. They are distributed on drug manufacturers' websites or provided by sales representatives through doctors' offices. The coupons reduce patients' copayments when they fill a prescription at pharmacies. Suppose one-month supply of a branded drug costs \$150 and its generic alternative costs \$30. A patient's insurance copayments for the two drugs are \$40 and \$10, respectively. Without a copay coupon, the patient, who is indifferent between the branded drug and generic alternative, would choose the less expensive generic and the insurer pays $20 for the prescription. If the branded drug manufacturer gives a copay coupon that reduces the out-of-pocket cost to $5, the patient would choose the branded drug and the insurer must pay $110. In this case, the branded drug manufacturer helps the patient to pay $35 for the copayment and earns $110 from the insurer.

Many branded drug manufacturers started to offer copay coupons as their drugs were losing patent protection in recent years. In December 2010, Pfizer launched a "Lipitor for You" program which allowed patients to pay as little as $4 for a month's supply of Lipitor, the best-selling drug in the history of pharmaceuticals. One month's supply of Lipitor normally has a retail price of $150 and the copay for generic Lipitor is about $10. Thus, the $4 copay program offered by Pfizer was very attractive, helping to keep about one-third of Lipitor's prescriptions within five months of its patent expiration in late 2011.[3] Many top-selling drugs followed the strategy as they were coming off patent, including blood thinner Plavix and blood pressure drug Diovan. Spending on copay coupons in 2011 is estimated to be $4 billion.[4] This is close to the aggregate spending on direct-to-consumer advertising ($4.3 billion) and accounts for two percent of gross branded drug sales in U.S.

Copay coupons can help to combat generic entry by lowering the costs for patients and influencing doctors' decisions. In March 2013, 53.5% of the 374 copay coupons found from www.internetdrugcoupons.com, a large drug coupon website, are for branded drugs with generic alternatives.[5] As many blockbuster drugs are going off-patent by 2015 and few new drugs are available to replace the revenue lost from patent expiration, branded drug companies have worked hard to retain their

[3] "New Coupons Aim To Keep People Off Generic Drugs," Associated Press, August 20, 2012.

[4] `http://www.pharmexec.com/pharmexec/article/articleDetail.jsp?id=755091`

[5] Among the copay coupon programs, 53.5% are from drugs with within-class generic alternatives, 8.3% are from drugs with FDA-approved therapeutic equivalents, and 38.2% are from drugs without lower cost alternatives. Source: Ross, J. S., and A. Kesselheim (2013): "Prescription-Drug Coupons - No Such Thing as a Free Lunch," *The New England Journal of Medicine*, 369, 1188-1189.

revenue after generics enter the market. Copay coupons can help to price compete with generics without cutting the full branded prices. Using copay coupons, branded drugs need to pay part of the out-of-pocket cost for patients. Since patients' share is usually less than one-third of full drug price, the benefit from insurance payments usually exceeds the cost of copay coupons. In addition, a report by Credit Lyonnais Securities Asia (CLSA) revealed that that 80% of physicians polled were more likely to prescribe a drug with a copay coupon.[6] Therefore, copay coupons could effectively influence doctors' decision to use brand-name drugs over generics by paying part of the copayment for patients.

The "shadow claims system" of copay coupons further contributes to their popularity. In the shadow system, prescription information is first sent to a pharmacy benefit manager (PBM), who processes prescription drug claims for employers, for adjudication. After the PBM adjudicates the prescription and sends the copay information back to the pharmacy, copay coupon programs reduce the copay for the coupon user. Thus, copay coupon programs are invisible to PBM's or insurers.[7] Because of the invisibility of copay coupon programs, insurers could not reject the use of copay coupons unless they remove the drug completely from their prescription drug list.

Copay coupons are banned in federal health programs, including Medicaid and Medicare, because they are considered illegal kickbacks that encourage unnecessary spending.[8] In the commercial market, Massachusetts had been the only state that prohibited copay coupons. However, in 2012 the Massachusetts government loos-

[6] http://www.uhc.com/pharmacy/news_and_updates/drug_copay_coupons.htm

[7] "How Copay Coupons Could Raise Prescription Drug Costs by $32 Billion Over the Next Decade," Pharmaceutical Care Management Association, November 2011.

[8] The prohibition does not apply to the insurance sold through the online health insurance marketplaces which began on October 1, 2013. The United States Department of Health and Human Services held that the insurance offered through the exchanges is not federal health care program subject to the prohibition.

ened the restriction by allowing manufacturers of branded drugs without competing generic equivalents to offer coupons.

1.3 Data

Data are obtained from four sources. First, data on pharmaceutical sales are provided by IMS Health. In the data, I observe national retail dollars and unit sales of each molecule/form/strength combination at monthly frequency from January 2003 to August 2011. Drug manufacturers and whether a drug is branded or generic are also provided in the data. Second, physician advertising data from Encuity Research contain monthly product level spending on detailing, medical journal advertising, and events and meetings.[9] Direct-to-consumer advertising data are obtained from Ad$pender database from Kantar Media. In the data set, they have monthly advertising spending for each product/media/market combination from January 2003 to August 2011. The media include television, radio, magazines, newspaper, internet and outdoor. There are national advertising as well as local advertising. For most of the drugs in the research, the spending is concentrated on national advertising. Thus, local advertising is ignored. Finally, copayment data are obtained from the MarketScan Research Databases through National Bureau of Economic Research (NBER). The files have prescription level claim data on copayments and full prices from about 150 employers, covering 40 million enrollees in the United States.[10]

I focus on the markets of HMG-CoA reductase inhibitors (statins), the major cholesterol medicines or lipid regulators. There are several reasons to look at this market for research on copay coupons. First, the cholesterol drug market is large. Cholesterol drugs were the third largest therapeutic class by spending in 2011, at

[9] Free sample data are available only after January 2007, so they are ignored.

[10] The data only cover privately insured individuals and do not include prescription drug claims from Medicare or Medicaid. The data limitation may lead to overestimated copayments for an average patient since generic utilization is higher in Medicare Part D plans.

20.1 billion US dollars. There were over 260 million prescriptions filled in 2011 and nearly 20 million Americans regularly used a cholesterol medicine.[11] Second, cholesterol medicines ranked first in spending on direct-to-consumer advertising among all therapeutic classes in years 2009 to 2011. Cholesterol drug manufacturers together spent on average 500 million dollars each year on DTCA.[12] This is evidence that firms in the cholesterol drug market care a lot about communication with consumers. Finally, entry of generics further intensified competition and adds to variation in pricing and advertising.[13] Two of the ten statins lost patent protection during the sample period and their generic versions entered right after their patent expiration. Entry of generic drugs dramatically changes the competitive environment, creating an opportunity to learn how consumers switch from branded drugs to generics.

The statins in the sample and their relevant facts are summarized in Table 1.1. The variations in class, molecule, form, and whether generics are available serve as the basis for modeling substitution among drugs. There are two classes: statins and statin combinations. Statins entered the market early in 1987 and statin combinations are relatively new as a treatment for high cholesterol. Statins combined with other molecules are treated as a different class since the combinations may have different effects on patients. Also, drugs in the classes come in three forms: tablet, sustained-action tablet, and capsule. Sustained-action is a mechanism that helps to dissolve a drug over time and release it more slowly and steadily into bloodstream so that a patient could take drugs less frequently. Because of the convenience from this mechanism, some consumers may prefer drugs in sustained-action tablet to drugs in tablet or capsule. Finally, three of the statins (Zocor, Pravachol and Mevacor) have

[11] *The Use of Medicines in the United States: Review of 2011*, IMS Institute.

[12] Source: Kantar Media A$Spender Database

[13] In this paper, entry is treated as exogenous. Adding a model of entry would complicate the analysis because a dynamic model would be needed for entry and this is beyond the scope of the paper.

generic alternatives. They all have a maximum number of generic equivalents greater than ten, implying severe within-molecule competition after patents expire. I treat the generics from different firms as separate products in the model for demand and supply.[14]

Note that Pfizer launched the "Lipitor for You" program and started to distribute copay coupons in December 2010. There are nine months of sales data in which some prescriptions for Lipitor may be associated with copay coupons. Since coupon use is not observable in the data, I do not consider the impacts of Lipitor's copay coupons in the estimation.

1.3.1 Sales and Prices

Figure 1.1 presents sales measured by patient-days. To make drugs with different strengths comparable in sales, I transform unit sales into patient-days by dividing the total number of milligrams sold by the recommended daily dosage.[15] The market grows over time with introduction of new drugs as well as entry of generics. The three major drugs are atorvastatin (Lipitor), simvastatin (Zocor), and rosuvastatin (Crestor). Zocor loses its patent protection in mid 2006. Within half a year after the patent expires, generic simvastatin quickly takes over the market of Zocor and expands the overall statin share. Atorvastatin (Lipitor) dominates the market until late 2006 when many generic simvastatins are available. The steady growth in the market share of Crestor, rolled out in 2003, also contributes to drops in the other branded drugs' sales.

[14] Some of the generics are "authorized generics," which are approved by FDA as brand-name drugs but marketed as generic drugs. According to a FTC report on authorized generics, pricing decisions by outside licensees typically are independent of the brand. The report also provides evidence of competition between authorized generics and their branded counterpart. In my data, all authorized generics are independent licensees. Thus, I treat them as competitors to the branded drug with the same molecule.

[15] The daily recommended dosage data are obtained from Clinical Pharmacology, an online database for drug information widely used by hospitals and retail pharmacies in US.

Figure 1.2 shows prices per patient-day. All prices are adjusted to January 2003 dollars. Most branded drugs have a price between one and three dollars per patient-day. Prices are quite stable for drugs under patent protection. In contrast, drugs with generic alternatives experience some price changes. Prices of branded simvastatin (Zocor), pravastatin (Pravachol) and lovastatin (Mevacor) tend to fall slightly as generics just enter but they move back to the original level when generic competition intensifies and generic prices become very low. The pricing pattern suggests that branded drugs would price compete with generics for the general consumers when there are only a few generics available. As many generics enter the markets, they choose to concentrate on the consumers with a strong preference for branded products.

Table 1.2 presents the summary statistics for the full sample as well as the subsamples for branded drugs both with and without generic equivalents, and generics. Each observation is a combination of month, molecule, brandedness, form, and manufacturer. Generics from different manufacturers are treated separately. On average, branded drugs without generic equivalents have the highest price and largest market share. The market share of branded drugs becomes much lower after generic entry while their prices are only slightly lower than before generics enter. Generic prices are on average one fourth of branded prices. An average generic accounts for 0.5% of market, compared to 2% for an average branded drug under patent protection. This shows that even if there are many branded drugs in the markets, patents still protect drugs from severe competition. After a patent expires, a large number of generic entrants simply take over the market from branded drugs.

1.3.2 Advertising

Figure 1.3 demonstrates the patterns of advertising to physicians and Figure 1.4 for direct-to-consumer advertising. Spending on advertising to physicians is the aggregate spending on detailing, journal advertising, and events and conferences. Direct-to-consumer advertising spending is the sum over national advertising spending through various channels. Also, only the four most advertised branded drugs are included because the advertising spending by the other products is relatively small. First note that physician advertising and DTCA spending for branded rosuvastatin is very large in the first few months of rollout to inform doctors and consumers of the existence of the new drug. Second, branded simvastatin and pravastatin start to cut advertising spending as their patent expiration dates approach. Their physician advertising spending drops to a very low level after generics become available. The drop in DTCA spending happens about one year before their patents expire. Table 1.2 shows that on average branded drugs' advertising spending is only 6% of the advertising spending before generics enter. These patterns imply that branded drugs have little incentive to invest in advertising after patent expires since consumers and physicians encouraged by advertising to use the drugs may choose the less expensive generic versions. Also, the patterns are consistent with the strategic investment story discussed in Ellison and Ellison (2007). Branded drug manufacturers have an incentive to reduce advertising to make a market less attractive to generic entrants although it is not clear that the strategy is effective in this case.

1.4 Model

In this section, I discuss the models of demand, supply, and copayments. I consider a random-coefficient discrete-choice demand model. The error structure is based on the model used in Bresnahan et al. (1997), which allows unobserved preferences to be correlated across multiple nests. By differentiating products along multiple dimensions, we can capture rich substitution patterns among drugs. To capture consumer heterogeneity in their preference for branded drugs, I use a simple two-type version of the random-coefficient model following Berry et al. (2006). Also, I construct a static supply side model to estimate marginal costs, identify some parameters for consumer heterogeneity, and to more precisely estimate the parameters from demand. Finally, the model for copayments builds the relationship between full price and out-of-pocket costs for consumers.

Before discussing the model set-up, it is worth being clear about a couple of limitations of the current analysis. First, like most of the literature on pharmaceutical demand, I assume that doctors and patients jointly make decisions to maximize utility. Thompson (1993) discusses that conflicts of interest between physicians and patients could arise from gifts given by drug companies to physicians, physicians' risk sharing in health maintenance organizations and hospitals, research on patients, etc. In recent years, government intervention and self-regulation by pharmaceutical industry have aimed to alleviate the problem and the conflicts of interest could be less serious.[16] Second, I do not try to consider how the use of coupons affects the entry of either branded or generic drugs. This could be addressed by adding an endogenous entry model to the current framework, but I leave this type of question to future work.

1.4.1 Demand

In the demand model, a consumer makes a discrete choice given a set of product characteristics. The consumer here refers to the combination of patients and doctors. I assume that they make joint decisions to maximize utility and ignore the possible principal-agent problem.[17] A product here is a combination of molecule, form,

[16] For example, The Physician Payments Sunshine Act, effective on August 1, 2013, requires manufacturers of drugs that participate in U.S. federal health care programs to report certain payments and items of value given to physicians. In 2008, Pharmaceutical Research and Manufacturers of America (PhRMA) strengthened the Code on Interactions with Healthcare Professionals to ensure that biopharmaceutical marketing practices comply with the highest ethical and professional standards.

[17] Lack of micro-level data on prescribing prohibits me from distinguishing the two roles in decision making.

brandedness, and firm. There are four dimensions along which products are differentiated: classes, branded/generic, molecules, and forms. As shown in Table 1.1, drugs of different molecules can be classified into statins and statin combinations. Three molecules have generic versions and two molecules have multiple forms.

An individual in period t chooses from J_t products, indexed $j = 1, 2, ..., J_t$. The indirect utility consumer i obtains from j in period t is

$$u_{ijt} = \alpha_i p^c_{jt} + x'_{jt}\beta + \mu_j + \xi_{jt} + \epsilon_{ijt}, \tag{1.1}$$

where p^c_{jt} is the copay for product j in time t, and x_{jt} is a set of time-varying observed product characteristics. μ_j is product fixed effect and ξ_{jt} is a time-varying component that captures unobserved demand shocks. Idiosyncratic taste parameter, ϵ_{ijt}, is assumed to be independent across consumers but correlated among products. The mean utility for product j in time t is $\delta_{jt} = x'_{jt}\beta + \mu_j + \xi_{jt}$. Consumers have an outside option, which includes non-drug treatments and no treatment. I normalize the utility of the consumer from this outside option to zero because I cannot identify relative utility levels.

The vector x_{jt} has several time-varying components that may affect consumer utility. First, I include the logarithm of physician advertising spending by product j in time t, $\log(1 + AD_{jt})$, and the logarithm of DTCA spending, $\log(1 + AC_{jt})$. Allowing advertising spending to enter directly to consumer utility, I assume that the two types of advertising have persuasive effects.[18] Second, to consider the spillover effects of advertising, I add physician advertising and DTCA spending from the other drugs with the same molecule, $\log(1 + ADOT_{jt})$ and $\log(1 + ACOT_{jt})$. These two variables help explain the branded drugs' advertising pattern close to patent expiration date, since free-riding of generics on branded advertising would reduce

[18] While literature on DTCA suggests that DTCA is more informative than persuasive, allowing the informative role of DTCA can complicate the model and largely increase computational burden. The DTCA here can be better viewed as a variable to control for time-varying product characteristics.

the incentive to invest in advertising. Third, I include time dummies for each period and time-since-entry dummies for each of the first twenty four months after drug entry. Time dummies capture the change in the quality of outside goods and time-since-entry dummies handle increasing consumer awareness of the existence of a new drug.

I model two types of consumers who differ in their choice sets and price sensitivities. High type consumers are assumed to have a strong preference for branded drugs and they consider only branded drugs. Low type consumers make a choice among all drugs available.[19] The price coefficient α_i is

$$\alpha_i = \begin{cases} \alpha^H & \text{if i is high type} \\ \alpha^L & \text{if i is low type} \end{cases}. \tag{1.2}$$

The consumer heterogeneity helps explain price differentials between branded and generic drugs. In the absence of consumer heterogeneity, the high price of branded drugs after patent expiration would be captured by a jump in marginal cost, which is not consistent with intuition. Moreover, estimating two types of consumers enables me to explore how coupon targeting affects welfare changes from copay coupons in the counterfactual in which coupon users are all low type.

To allow ϵ_{ijt} to be correlated among products, I follow McFadden et al. (1978) and assume the unobserved idiosyncratic parameter has a generalized extreme value (GEV) distribution with multivariate cumulative distribution function

$$F(\epsilon_{i0t}, \epsilon_{i1t}, ..., \epsilon_{iJt}) = \exp\left[-G\left(e^{\epsilon_{i0t}}, e^{\epsilon_{i1t}}, ..., e^{\epsilon_{iJt}}\right)\right],$$

which implies that the market share of product j in time t from the high type is

[19] The assumption on the consumer heterogeneity in choice sets is equivalent to including a dummy variable in high type's indirect utility and forcing the coefficient on the dummy variable to be negative infinity.

given by

$$s^H_{jt} = \frac{e^{\delta_{jt}+\alpha^H p^c_{jt}} G^H_j \left(e^{\delta_{i0t}}, e^{\delta_{i1t}+\alpha^H p^c_{1t}}, ..., e^{\delta_{iJt}+\alpha^H p^c_{Jt}}\right)}{G^H \left(e^{\delta_{i0t}}, e^{\delta_{i1t}+\alpha^H p^c_{1t}}, ..., e^{\delta_{iJt}+\alpha^H p^c_{Jt}}\right)}, \tag{1.3}$$

where G^H_j is the partial derivative of G^H with respect to the j^{th} argument. Similarly the market share of product j in time t from the low type is given by

$$s^L_{jt} = \frac{e^{\delta_{jt}+\alpha^L p^c_{jt}} G^L_j \left(e^{\delta_{i0t}}, e^{\delta_{i1t}+\alpha^L p^c_{1t}}, ..., e^{\delta_{iJt}+\alpha^L p^c_{Jt}}\right)}{G^L \left(e^{\delta_{i0t}}, e^{\delta_{i1t}+\alpha^L p^c_{1t}}, ..., e^{\delta_{iJt}+\alpha^L p^c_{Jt}}\right)}. \tag{1.4}$$

Following Bresnahan et al. (1997), I specify G^H and G^L as

$$G^H \left(e^{\delta_{i0t}}, e^{\delta_{i1t}+\alpha^H p^c_{1t}}, ..., e^{\delta_{iJt}+\alpha^H p^c_{Jt}}\right) = e^{\delta_{0t}} + \sum_l a_l \left[\sum_k \left(\sum_j I(j,k,l) BRAND_j e^{\frac{\delta_{jt}+\alpha^H p^c_{jt}}{\rho_l}} \right)^{\rho_l} \right] \tag{1.5}$$

$$G^L \left(e^{\delta_{i0t}}, e^{\delta_{i1t}+\alpha^L p^c_{1t}}, ..., e^{\delta_{iJt}+\alpha^L p^c_{Jt}}\right) = e^{\delta_{0t}} + \sum_l a_l \left[\sum_k \left(\sum_j I(j,k,l) e^{\frac{\delta_{jt}+\alpha^L p^c_{jt}}{\rho_l}} \right)^{\rho_l} \right], \tag{1.6}$$

where $I(j,k,l)$ is an indicator variable taking on the value of one if product j has the k^{th} value of the l^{tj} characteristic and ρ_l is the nesting parameter along the l^{th} dimension. $BRAND_j$ is an indicator equal to one if j is a branded drug and zero otherwise. The scaling parameter a_l is defined as

$$a_l = \frac{1-\rho_l}{\sum_{l=1}^{L}(1-\rho_l)}.$$

The market share for product j in time t is the weighted average of the market share from the two types. Assume high type consumers account for λ_t fraction of the market in time t. The market share for product j in time t can be expressed as

$$s_{jt} = \lambda_t s^H_{jt} + (1-\lambda_t) s^L_{jt}. \tag{1.7}$$

1.4.2 Firm Behavior

Assume there are $f = 1, 2, ..., F_t$ firms in period t competing in a Bertrand-Nash game. Firm f produces a subset of J products, J_f. The profit for firm f, omitting the time subscript, is

$$\Pi_f = \sum_{j \in J_f} (p_j - mc_j) M s_j(p, AC, AD, \xi; \theta) - AD_j - AC_j, \tag{1.8}$$

where s_j is the market share of product j, mc_j is the marginal cost of product j, and M is the market size. AD_j and AC_j are the spending on advertising to consumers and physicians, respectively. The marginal cost is assumed to be

$$\log(mc_{jt}) = \eta_j + g(t) + h(\tau_{jt}) + \omega_{jt},$$

where η_j is the product fixed effect, $g(t)$ the function for time trend, $h(\tau_{jt})$ the function for time since entry, and ω_j is the (time-varying) unobserved cost shocks.

Given the prices, product attributes, advertising spending and marginal costs, firms simultaneously choose prices to maximize profits.[20] The first order condition with respect to price is given by

$$s_j + \sum_{k \in J_f} (p_k - mc_k) \frac{\partial s_k}{\partial p_j} = 0. \tag{1.9}$$

1.4.3 Copayments

The price faced by insured consumers is a small share of the full price. In the United States, insurance plan enrollees may pay a fixed amount for each prescription regardless of the drug cost (copayment), or a percent of the prescription drug cost (coinsurance). The tier pricing system designed by insurance companies usually puts

[20] I do not consider the decision on advertising in firms' problem and assume that there is no strategic interaction in advertising. In the counterfactuals later, I do not solve for new advertising levels either.

less expensive drugs in lower tiers and requires a smaller copayment or coinsurance from enrollees. For example, a typical three-tier system has generics in tier one, branded drugs without generic substitutes in tier two, and branded drugs with generic substitutes in tier three. According to the 2011 Annual Survey by Kaiser Family Foundation, 72% to 85% of covered workers have copays for drugs listed in the first three tiers and 7% to 11% of covered workers have coinsurance.[21]

To have a model that nests copayment and coinsurance for an average consumer, I assume

$$\log(p^c_{jt}) = \gamma_0 + \gamma_1 \log(p_{jt}), \tag{1.10}$$

where p^c_{jt} is the cost shared by the consumer for product j in period t and p_{jt} is the full price for product j in period t. The first term (γ_0) is the fixed amount and the second term $\gamma_1 \log(p_{jt})$ is the cost as a part of the full price. The model thus incorporates the two types of cost-sharing systems. In addition, the log-log model would be able to accommodate the fact that a high full price charged by drug manufacturers would not be proportionally passed on to enrollees. The marginal growth of their costs will be diminishing for $\gamma_1 \in (0, 1)$. Similarly, a very low full price would not make enrollees' out-of-pocket costs proportionally lower since they are responsible for a basic payment for each prescription. As illustrated in Figure 1.5, the log-log model predicts the copayment for consumers by pushing down a high full price and moving up a low full price.

1.5 Estimation

The estimation of demand parameters closely follows Berry et al. (1995, 2004) and Nevo (2000). I assume that the demand and pricing unobservables are mean independent of a set of instruments at the true parameters. That is, $E\left[\xi_j(\Theta_0) \mid Z\right] =$

[21] http://kff.org/health-costs/report/employer-health-benefits-annual-survey-archives/

$E\left[\omega_j(\Theta_0) \mid Z\right] = 0$. Using the contraction mapping suggested by BLP, I am able to compute ξ_j given a set of parameter values and observed market shares:[22]

$$\xi_{jt} = \delta_{jt}(s,\theta) - x'_{jt}\beta - \mu_j. \tag{1.11}$$

The marginal cost is computed from the first order condition:

$$mc = p - \Delta(\theta,\delta)^{-1}s(\theta,\delta) \tag{1.12}$$

where $\Delta_{j,k} = -\partial s_k/\partial p_j I_j$ with I_j equal to one if j and k are produced by the same firm. Then we can derive

$$\omega = log(p - \Delta(\theta,\delta)^{-1}s(\theta,\delta)) - \eta - g - h(\tau). \tag{1.13}$$

Estimation of the parameters is undertaken by the generalized method of moments (GMM). I minimize the objective function of $\Lambda' Z W Z' \Lambda$, where W is the weighting matrix. Let Z_ξ be the instruments for ξ and Z_ω be the instruments for ω. The sample moments are (the time subscript are suppressed)

$$Z'\Lambda = \left[\begin{array}{c} \frac{1}{J}\sum_j Z_{\xi,j}\xi_j(\alpha,\beta^{AC},\beta^{AD}) \\ \frac{1}{J}\sum_j Z_{\omega,j}\omega_j(\alpha,\beta^{AC},\beta^{AD},\eta_j) \end{array} \right]. \tag{1.14}$$

The choice of instruments for price and advertising relies on the identifying assumption used in Bresnahan (1987) and Berry et al. (1995). I assume the location of each drug in product space is exogenous and a drug's markup, which is a function of prices and advertising levels, is correlated with its relative isolation in the product space. Since I do not include product characteristics in the indirect utility, rather than summing up the characteristics of own products and the other firms' products as in BLP, I follow Arcidiacono et al. (2013) and count the number of products in a

[22] Because of the model specifications, the contraction mapping here is slightly modified and the proof of invertibility is put in the appendix.

category defined in various ways. Specifically, I use the number of molecules for the same form, number of molecules of the same form in the same class, whether generics are present in the same form, whether generics are present in the same molecule, number of generics present of the same molecule, number of generics present of the same form, and number of generics present of the same form in the same class.[23]

I discuss identification in an intuitive way. Nesting parameters (ρ_l's) are identified from changes in aggregate market share for each nest when the number of products in a nest varies. Take as an example the case of a nested logit model with molecules as nests. If the nesting parameter is one, the model reduces to a simple logit model. The market share would be roughly the same for each drug if they share similar product characteristics. If the nesting parameter is zero, drugs of the same molecule are perfect substitutes. Adding one drug to a molecule nest or changing the price of a drug in a nest does not affect the market shares of drugs in the other nests.

Identification of the fraction of high type consumers (λ) relies on the first order conditions with respect to prices on the supply side. If λ is zero, more competition from generics will drive branded drugs' prices down. If λ is one, there will be zero market share for generics and branded drugs' prices will not change as generics enter. Thus, firms' pricing decisions help to identify the fraction of high type consumers. Once λ is identified, the price coefficients for each type (α^H and α^L) could be identified from variations in prices of both branded and generic drugs. Identification of the linear parameters is more straightforward. Variations in advertising spending across products and over time helps to identify their coefficients in the demand model.

[23] The F-statistics based on usual standard errors from the first-stage regressions of endogenous variables on the instruments range between 29 and 113, suggesting that the instruments are relevant.

1.5.1 Results

Table 1.3 shows the results for copay estimation from 2003 to 2009. The estimates for each year are the weighted average over plans based on the empirical distribution of the number of plan enrollees. The estimates for the two parameters are similar over the years. They are all significant at 5% level and the large sample size makes the standard errors small. Those estimates are used to predict the copayments for demand estimation and counterfactuals. For the years after 2009, I use the estimates for 2009 to predict the copayments.

Table 1.4 presents the estimates and standard errors for nesting parameters and the most important linear parameters in demand. Most estimates are statistically significant at 5% level. The nonlinear parameter estimates reveal how consumers substitute among the products. The estimated nesting parameter for molecule is about 0.56 while the other nesting parameter estimates fall between 0.63 and 0.67. This means that within-molecule substitution is the strongest, which is consistent with the sudden changes in the market share of branded drugs when their generic alternatives become avaiable. The estimate for the fraction of high type consumers is about 14%. The price coefficient for high type consumers is estimated to be -8.7, compared to -27.2 for low type consumers. Obviously, a non-trivial portion of consumers has a strong preference for branded products and they care much less about the drug copays than do low type consumers.

The large price coefficients do not mean that we should expect price elasticities to be extremely large. Use of copay rather than full price in the demand model generates larger price coefficients since copays are a small share of full price. According to the copay estimates, one dollar increase in full price results in an average of \$0.21 increase in branded copay and \$0.26 increase in generic copay. This implies that insurers would pass only one fourth of drug cost increase on to consumers and thus the price elasticities shown later would generally be small.

Linear parameter estimates show the effects of advertising and increasing consumer awareness since a drug's rollout. Recall that I include in the indirect utility the logarithm of advertising spending. Therefore, in interpreting the estimates for advertising parameters, we need to control for the advertising levels. The estimate for $\log(1 + AC)$ is larger than the estimate for $\log(1 + AD)$, which implies that, at the same level of spending, DTCA is more effective than physician advertising in this market. Moreover, the estimate for $\log(1 + ADOT)$ is larger than the estimate for $\log(1 + AD)$, and the estimate for $\log(1 + ACOT)$ is also larger than the estimate for $\log(1 + AC)$. The strong spillover effects of physician advertising and DTCA are probably resulted from the small investment in advertising from branded drugs after generic entry and the increase in generic market share at the same time. The estimates for $\log(1 + ADOT)$ and $\log(1 + ACOT)$ capture generics' large gain from advertising by their branded rivals. In addition, the time-since-entry coefficient estimates suggest on average it takes about two months for a new drug to get attention. The first two estimates for the time-since-entry dummies are not significant at the 5% level. As a drug is on the market for more than two months, the effect of time-since-entry gets stronger and more significant.

Table 1.5 shows the selected estimated cost parameters. The results include the fixed effect estimates for major branded drugs and top-selling generic drugs. Also, the table presents the estimated parameters on the dummies for one month, one year and two years since entry. Other things being equal, branded atorvastatin, rosuvastatin, and lovastatin have similar marginal costs to each other while branded pravastatin has a higher marginal cost. Branded drugs' marginal costs are 2.2 to 3.6 as large as those of their generic alternatives. Moreover, comparing the estimated coefficients on the dummies for time since entry, I find that marginal cost is declining quickly since a drug's entry. This implies that drug production gets more efficient as a drug is on the market longer.

Table 1.6 contains price elasticities based on data for June 2006. At the end of this month, the patent of simvastatin (Zocor) expires and generic simvastatin start to enter. Table 1.7 summarizes the price elasticities for September 2006, three months after generic simvastatins' entry. The price elasticities for these two months help to understand the results of counterfactuals, discussed in detail later, in which Zocor aims to compete with generics using copay coupons. I discuss the results for June 2006 first. The columns are the percentage change in market share in response to one percent increase in the price of drugs in the rows. For example, the market share of ezetimibe/simvastatin (tab) is predicted to increase by 0.20% if amlodipine/atorvastatin (tab) raises price by 1%. Results for all branded drugs and the top-selling generic drugs are reported. There are several interesting findings from the table. First, all own price elasticities are larger than one in absolute value. Under the assumption of profit maximization, an own price elasticity greater than one in absolute value implies positive marginal revenue. If the own price elasticities were less than one in absolute value, then the marginal revenue and the implied marginal cost would be negative, which would make it difficult to construct a model for the counterfactual analysis. Second, own price elasticities of generic pravastatin and simvastatin are especially large. Recall that only low type consumers would choose generics and they are more price sensitive than high type consumers. Thus, a price change in generics would result in a larger change in their own market share. A mix of high and low type consumers for branded drugs leads to a lower own price elasticity.

Third, the cross price elasticities of drugs with similar characteristics are generally larger than the other cross price elasticities. For example, one percent price decrease of branded simvastatin in tablet (Zocor) has the largest impact on generic simvastatin, leading to a 3% decrease in share. Among branded drugs, those of the same form as Zocor are more affected in general. A one percent decrease in price of Zocor would lower the share of lovastatin in tablet by 1.2% while the price change only results in 0.3% decrease in the share of lovastatin in sustained-action tablet.

In addition, the substitution between branded simvastatin and generic simvastatin is quite asymmetric. The price impact of branded simvastatin on generic simvastatin is much larger than the impact of generic simvastatin on branded simvastatin. This is primarily because the market share difference between branded simvastatin and generic simvastatin is large when the generic just enter. In this month, branded simvastatin has about 5% of market share while the generic simvastatin only accounts for 0.26% of the market. The price change of branded simvastatin, therefore, has a larger impact on the percentage change of the generic's market share than the other way around.

Finally, all results in Table 1.7 are similar to those of Table 1.6, except for the elasticities with respect to the price changes of branded simvastatin. In September 2006, the market share of branded simvastatin decreases to 0.93% as generic simvastatins take over its market. The changes in the market share of generics given a price change of branded simvastatin are almost zero since most of the consumers of branded simvastatins in this month probably have a strong preference for branded drugs and they would not switch to generics. This implies that the price changes by branded simvastatin would affect the other branded drugs more than the generics when generic simvastatins are popular.

1.6 Counterfactual Analysis

Using the estimates for demand and supply parameters, I construct counterfactuals to explore the effects of copay coupons on social welfare. I assume that Merck, the manufacturer of branded simvastatin (Zocor), decides to issue copay coupons to consumers when its patent expires. I choose Zocor as the issuer of copay coupons for two reasons. First, I observe the entry of generic simvastatins in the data and from estimation results I learn how consumers switch to those generics from the other drugs. Second, Zocor is the best-selling branded drug before Lipitor. Understanding the outcomes of Zocor's copay coupon program would shed some light on the welfare impacts of Lipitor's coupon program. It is thus interesting to experiment with Zocor to study the welfare implications of the new pricing tactic.

There are some simplifying assumptions that facilitate the counterfactuals. The fraction of consumers receiving copay coupons is assumed to be exogenous since the manufacturer could not fully control how many consumers actually receive the coupons. Also, the insurance copay is still determined by the copayment model. Given a full price, insurers would charge consumers a share of drug costs using the estimated copay formula and they would not change the formula. Furthermore, I only consider the effects of issuing coupons during the first five months (July 2006 to Nov 2006) after Zocor's patent expires. During this period, there are three manufacturers for generic simvastatins: Teva, Ranbaxy, and Dr. Reddy's Laboratories. Teva and Ranbaxy are the first challengers of Zocor's primary U.S. patent and were granted 180-day exclusivity by FDA to sell generic simvastatins. Dr. Reddy's Laboratories received a license from Zocor to sell authorized generic simvastatins.[24] No other firms are allowed to enter the market of simvastatins during these five months.[25] There are two main reasons to restrict the experiment to the 180-day exclusivity period. First, with limited generic competition, branded drug manufacturers have more incentive to use copay coupons than when generic competition is severe. Once there are many generics in the market, generic prices are usually very low and the coupons would

[24] According the FTC report on authorized generics, outside licensees' pricing decisions are typically independent of the branded drug manufacturers. Thus, I treat Dr. Reddy's Laboratories as a competitor to Merck rather than a partner.

[25] In late December 2006, other generics started to enter the market and thus I exclude this month in the policy simulation.

not be as attractive to consumers. Second, the lack of a model for entry makes it difficult to tackle the possible generic entrants after the exclusivity period.

Suppose there are J_t drugs in period t and the fraction of consumers receiving a copay coupon from Zocor is $q \in (0,1)$, firms compete in a Bertrand game and maximize their own profits, under the full information assumption, by simultaneously choosing their full prices $(p_1, p_2, ..., p_{J_t})$ and, for Zocor, the copay with a coupon $(\tilde{p}^c_Z)$.[26] Insurers decide the insurance copay for each drug $(p^c_1, p^c_2, ..., p^c_{J_t})$ using the copay formula. Consumers with a copay coupon face copays $[p^c_1, p^c_2, ..., \min\{p^c_Z, \tilde{p}^c_Z\}, ..., p^c_{J_t}]$ and consumers without a copay coupon face copays $[p^c_1, p^c_2, ..., p^c_Z, ..., p^c_{J_t}]$. Note that a consumer with a copay coupon from Zocor will compare the copay using the coupon with the insurance copay. If the copay from coupon is higher than the insurance copay for Zocor, the consumer will not use the coupon at all.

With the equilibrium prices, I can calculate the change in firm profits, consumer welfare and insurer's spending. The profit from Zocor in period t with copay coupons is

$$\begin{aligned}\pi_Z = M_t \left[(1-q)\left(p_{Zt} - mc_{Zt}\right) s^{NC}_{Zt} + q\left(p_{Zt} - p^c_{Zt} + \min\{\tilde{p}^c_{Zt}, p^c_{Zt}\} - mc_{Zt}\right) s^C_{Zt}\right] \\ - AD_{Zt} - AC_{Zt},\end{aligned}$$

where s^{NC}_{Zt} is Zocor's share from those without a coupon and s^C_{Zt} the share from those with a coupon. The revenue from those with a coupon is insurer's payment $(p_{Zt} - p^c_{Zt})$ plus the minimum of copay using a coupon and insurance copay $(\min\{\tilde{p}^c_{Zt}, p^c_{Zt}\})$. Under the GEV error structure, the expected consumer surplus in period t, measured in dollars, before copay coupons are used can be expressed as

$$E\left(CS_{0t}\right) = M_t \left[\lambda \left(\frac{1}{-\alpha^H}\right) \log(G^H_t) + (1-\lambda)\left(\frac{1}{-\alpha^L}\right) \log(G^L_t) + C_E\right], \qquad (1.15)$$

[26] In this case, choosing the copay for coupon users $(\tilde{p}^c_Z)$ is equivalent to choosing the coupon value $(p^c_Z - \tilde{p}^c_Z)$ since, given p_Z, p^c_Z is decided according to the copay formula and the formula is fixed throughout the game.

where C_E is the Euler's constant. The expected consumer surplus in period t, measured in dollars, after copay coupons are used is

$$E\left(CS_{1t}\right) = M_t\Bigg[\lambda\left(\frac{1}{-\alpha^H}\right)\left(q\log(G_t^{H,C}) + (1-q)\log(G_t^{H,NC})\right) + \\ (1-\lambda)\left(\frac{1}{-\alpha^L}\right)\left(q\log(G_t^{L,C}) + (1-q)\log(G^{L,NC})_t\right) + C_E\Bigg], \quad (1.16)$$

where $G_t^{H,C}$ and $G_t^{H,NC}$ are the G function values from high type consumers with and without a copay coupon, respectively. $G_t^{L,C}$ and $G_t^{L,NC}$ are the G function values from low type consumers with and without a copay coupon, respectively. The change in the expected consumer surplus in period t is

$$\Delta E(CS_t) = E\left(CS_{1t}\right) - E\left(CS_{0t}\right). \quad (1.17)$$

The Euler's constants in $E\left(CS_{1t}\right)$ and $E\left(CS_{0t}\right)$ simply cancel out.

An additional assumption needs to be made to solve for equilibrium prices. Theoretically, Zocor would set $p_Z = \infty$ and $\tilde{p}_Z^c \approx 0$ for any $q > 0$. This way, Zocor can concentrate on the consumers with copay coupons and receive an infinite payment from the insurance companies for each of these consumers. Figure 1.6 illustrates how Merck's profit changes with Zocor's full price when $q = 0.05$, holding fixed the other prices and the copay with coupon. The profit from those who do not have a coupon reaches a maximum at the price of 2.4 while the profit from coupon users is strictly increasing in the full price of Zocor. Merck's total profit is always increasing in the full price of Zocor since the profit increase from coupon users is larger than the profit decrease from coupon nonusers beyond the price of 2.4. To limit the equilibrium prices to a reasonable level, I set an upper bound for Zocor's price equal to 1.25 times the current Zocor's price. Beyond that point, I assume that the insurer will remove the drug from its prescription drug list. The choice of the upper bound is based on the fact that Lipitor's real price increased by about 23.8% over the two years since its coupon program was introduced in late 2010.[27] The price increase by Lipior has led some insurers and pharmacy benefit managers to exclude Lipitor from their prescription drug lists.[28]

I consider two counterfactuals: (1) consumers who receive copay coupons and those who do not are equally price sensitive, and (2) consumers who receive copay coupons are all low type (on average, they are more price sensitive than those who do not receive coupons). In counterfactual 1, Merck cannot price discriminate based on price sensitivities since the composition of coupon users and nonusers is the same. Thus, the results from counterfactual 1 are all driven by the agency problem between insurers and patients. In counterfactual 2, I allow Merck to target coupons to the most price-sensitive individuals to understand how coupon targeting and the agency issue together change the effects of coupons on welfare.

Within the two counterfactuals, I consider a case with no price response and a case with price response. In the case with no price response, Merck sets the copay for coupon users to maximize its profit, holding all prices (including Zocor's price) fixed. In the other case, firms are allowed to respond by choosing prices to re-optimize profits. Comparing the results from the two cases could show how strategic interaction affects social welfare when copay coupons are introduced. In addition, I solve for equilibrium prices in each counterfactual for different degrees of coupon penetration to learn how expanding the coupon program affects welfare and profits.

1.6.1 Counterfactual 1: Baseline

Table 1.8 shows that Merck sets the copay for coupon users equal to less than half the insurance copay for Zocor when all full prices are fixed. The copay with coupons

---

[27] "Rx Price Watch Case Study: Efforts to Reduce the Impact of Generic Competition for Lipitor," American Association of Retired Persons, June 2013.

[28] For example, starting on January 1, 2013, UnitedHealthcare stops offering Lipitor to its plan enrollees because of Lipitor's high price and extensive use of copay coupons.

is also lower than the average insurance copay for generic simvastatins and close to the average insurance copay for the other generics. This results in an increase in Zocor's market share by more than three times when there are only 5% of consumers receiving copay coupons. As the fraction of consumers with a coupon increases to 50%, Zocor's market share becomes 37 times as large as its original level. The other drugs' combined market share drops by about 10 percentage points (40%). These show that Merck wants to aggressively lower the copay for coupon users to expand Zocor's market share.

The low copay set by Merck has a large impact on firm profits and insurance payments. Table 1.9 shows that consumers benefit from copay coupons because some of them could buy Zocor at a low price and those without coupons face the same price as before. Merck's profit from Zocor increases by 12 million dollars when only 5% of consumers receive coupons. The small fraction of consumers who get a coupon more than doubles Merck's profits from Zocor. At the same time, the other drugs' profits drop by about 5% and insurance spending grows by 15%. The net effect of the coupon program is negative mainly because of the large increase in insurance payment. The net effects of copay coupons on social welfare get more negative as the coupon program expands.

When firms are allowed to change prices in response to introduction of coupons, Merck would set a very low copay for coupon users and raise the full price of Zocor as high as possible. This way, Merck is able to attract most of the consumers with a coupon and earn a large profit from the insurer. Table 1.10 shows that all full prices of Zocor hit the upper bound (1.25 times the original price) and the copay for coupon users is lower than any insurance copay in the market. The pricing strategy implies that Merck views the market with coupons as its main source of profits. The low copays for coupon users make Zocor's market shares larger at any level of coupon penetration than its corresponding market shares under no price response.

On the other hand, the average equilibrium price of the other branded drugs becomes higher after coupons are introduced because Zocor's high prices mitigate price competition in the market. As coupon penetration rate increases, the average branded drug prices slightly go up. This is the result of different pricing strategies played by Merck and the other firms. As coupons become more popular, Merck sets a higher price for their branded drugs to make the coupons more attractive. Doing so would sacrifice some market share from those without a coupon. But given the large profits from Zocor's coupons, it is optimal to raise prices of the other branded products owned by Merck to increase Zocor's market share. The other firms, however, would want to cut the prices of their branded products to compete for coupon users as Zocor's coupon program gets larger. Obviously, the effect of higher prices set by Merck dominates the effect of lower prices set by the other drug manufacturers when the fraction of consumers with a coupon is less than or equal to 50%. When coupon penetration rates are larger than 55% (not shown in the table), the average branded prices start to drop since the higher prices of Merck's products make their market share smaller and this reduces their weights in the average price.

Table 1.11 shows that the coupon program increases consumer welfare, Zocor's profit, and insurer spending more than it does in the no-response case. The larger increase in consumer welfare comes from the lower copay for coupon users and the larger overall market share. Merck further improves its profits from Zocor by issuing coupons since they are able to raise the full prices in the price-response case. The gains of Merck and consumers come at the cost of insurance payments. The increase in insurer spending in the price-response case is 70% more than the increase in the no-response case. Insurance payments increase by 25% when 5% of consumers get a coupon and by 50% when the coupon penetration rate is 10%. In addition, when 5% of consumers receive coupons, the other branded drugs have a higher combined profit than before coupons are introduced. As coupons become more popular, change in the profits of the other branded drugs turns negative but the effects of coupons on the other branded drugs' profits are weaker than in the no-response case. When the coupon penetration is low, the high price of Zocor mitigates price competition among branded drugs in the market of those who do not have a coupon, which lessens the negative impacts of copay coupons on the profits from the other branded drugs.

The results from counterfactual 1 show that coupons benefit the issuer by taking advantage of the agency problem between insurers and patients. Copay coupons can help the issuer to induce consumers to buy its product by reducing their copays and earn a large profit from insurance companies. In most cases, the other firms are hurt by the coupon program because of a smaller market share. In the case with price response, insurer spending increases by 25% when only 5% of consumers receive a coupon. The net effect of copay coupons on social welfare is negative because the increase in insurance payments is always more than the gains of consumers and the coupon issuer.

1.6.2 Counterfactual 2: Coupon Targeting

In counterfactual 2, I assume Merck has the ability to direct coupons to the most price-sensitive types of consumers. Thus, all consumers who receive a coupon are assumed to be low type. Most of the results from counterfactual 2 are consistent with those of counterfactual 1. Thus, I focus on the major differences between the results in the counterfactuals and discuss how the ability to target coupons contributes to the differences.

First, in the no-response case of counterfactual 2, the copay for coupon users is slightly lower than in the no-response case of counterfactual 1. Table 1.12 shows that the copay is set to be \$7.45, compared with \$7.57 in counterfactual 1. Because now the coupon users are more price sensitive, the optimal copay for them becomes lower. Second, we learn from Table 1.13 that the increase in Merck's profits from

Zocor is larger than in counterfactual 1 and the decrease in the other branded drugs' profits is much smaller. In contrast, the loss in generics' profits becomes larger. In counterfactual 2, there are more high type consumers among those without a coupon, for a given coupon penetration rate, than in counterfactual 1. Since they only consider branded drugs, branded drugs' market shares and profits are higher than those of counterfactual 1. Also, generics are hurt more in counterfactual 2 because of a higher proportion of high-type consumers without a coupon who do not choose generics at all.

Third, Table 1.14 shows that Zocor's prices hit the upper bound again in the price-response case. The equilibrium copay for coupon users is slightly lower than the copay value in the price-response case in counterfactual 1 since in counterfactual 2 the coupon users are more price sensitive. Also, the other branded drugs' average optimal prices are higher than those in the price-response case of counterfactual 1. Since the other branded drug manufacturers know that those without a coupon are on average less price sensitive than those with a coupon, it is optimal for them to charge a higher full price to capture the consumers without coupons. It is interesting that the average optimal prices of the other branded drugs get higher as more and more low-type consumers have a coupon. This pricing pattern suggests that Merck's ability to target coupons to more price-sensitive consumers makes branded drug manufacturers care more about the consumers without coupons and exploit their lower price sensitivity by raising full prices.

Additionally, Table 1.15 shows that changes in the profits of the other branded drugs are all positive and the increase in the profits gets larger as more low-type consumers have a coupon. When coupons are introduced and targeted to low-type consumers, branded drug manufacturers are able to set a higher price and earn a larger profit than when coupons are randomly distributed. Zocor's targeting coupons to low-type consumers raises the proportion of high-type in the consumers without a coupon, which are the main source of profits for the other branded drugs. Thus, the average price sensitivity of consumers without a coupon is lower than in counterfactual 1 and this lets branded drug manufacturers set a higher price and earn more in counterfactual 2. The profits from Zocor are also larger than in counterfactual 1 because the difference between Zocor's price and the price of the other branded drugs narrows, which leads to a larger market share of Zocor in the consumers without coupons. Finally, consumer welfare gains do not change much compared to the gains in the price-response case in counterfactual 1. The effects of lower copay for coupon users and higher prices for coupon nonusers roughly cancel each other out.

To sum up, Zocor's coupon targeting helps screen consumers and lowers the average price sensitivity of those without a coupon. This further mitigates price competition among branded drugs and benefits not only Merck but the other branded drug manufacturers as well.

1.7 Conclusion

To understand the welfare implications of copay coupons, I estimate a model with consumer heterogeneity and rich substitution patterns and demonstrate the effects of copay coupons on the healthcare system in the counterfactuals. I consider two motivations to use copay coupons: the agency problem and coupon targeting. First, I find that the agency problem between insurers and patients can make copay coupons highly profitable. Copay coupons benefit the coupon issuer and consumers at the cost of larger insurance payments. The other firms are hurt in most cases since they could hardly compete for the consumers with a coupon. Second, when I allow for coupon targeting, the coupon benefits for the coupon issuer become larger because targeting enables coupons to screen consumers and mitigate price competition among branded drugs. In this case, most of the branded drug manufacturers also gain from the introduction of copay coupons.

There are several interesting extensions for the study of copay coupons in pharmaceuticals. First, copay coupons can be used for competition between branded drugs. In the paper, I only consider the case with a single coupon issuer after its patent expires. Competition using coupons when there are no generics available can expand the overall market and have different welfare implications. Second, solving for advertising spending with the introduction of coupons can improve the welfare analysis. Usually, branded drugs cut advertising spending after generics enter because of generics' free-riding problem. With copay coupons, branded drug manufacturers may have a stronger incentive to advertise since they would want to attract more coupon users. If advertising is valuable to consumers, the gain in consumer welfare from coupons could be larger with more advertising.

Finally, in addition to the agency problem and price discrimination, dynamics could motivate branded drug manufacturers to use copay coupons. For example, they may use coupons to build brand loyalty and help patients stay on their medicines after they stop offering coupons. Also, copay coupons could be used to deter generic entry since the low copay with coupons makes it very difficult for generics to compete with branded drugs. Incorporating these dynamic motivations would complement the static analysis in the paper and provide more insights on the welfare implications of copay coupons.

Table 1.1: Statins and Statin Combinations

Class	Molecule	Brand Name	Form	Brand Entry	1st Generic Entry	Max Num Generics
Statins	atorvastatin	Lipitor	TAB	Jan 1997	-	-
	fluvastatin	Lescol	CAP	Apr 1994	-	-
	fluvastatin	Lescol XL	SA TAB	Nov 2000	-	-
	lovastatin	Altoprev	SA TAB	Jul 2002	-	-
	lovastatin	Mevacor	TAB	Sep 1987	Feb 2002	11
	pravastatin	Pravachol	TAB	Nov 1991	May 2006	14
	rosuvastatin	Crestor	TAB	Aug 2003	-	
	simvastatin	Zocor	TAB	Jan 1992	Jun 2006	16
Statin Combinations	amlodipine/atorvastatin	Caduet	TAB	Mar 2004	-	-
	ezetimibe/simvastatin	Vytorin	TAB	Jul 2004	-	-
	lovastatin/niacin	Advicor	SA TAB	Dec 2001	-	-
	niacin/simvastatin	Simcor	SA TAB	Feb 2008	-	-

Table 1.2: Price, Market Share, and Advertising Spending

Variable	Obs	Mean	Std. Dev.	Min	Max
Full sample					
Price	3638	1.0642	1.0127	0.0378	5.0874
Share	3638	0.0082	0.0169	0.0000	0.0948
Physician ad	3638	1,544,225	4,112,393	0	30,083,698
DTCA	3638	709,144	2,894,856	0	31,339,548
Branded without generic equivalent					
Price	812	2.3090	0.7267	0.9864	5.0874
Share	812	0.0198	0.0267	0.0000	0.0948
Physician ad	812	6,741,131	6,380,512	0	30,083,698
DTCA	812	3,176,964	5,452,706	0	31,339,548
Branded with generic equivalent					
Price	335	2.2223	0.8930	0.6456	4.3290
Share	335	0.0010	0.0031	0.0000	0.0499
Physician ad	335	415,167	791,388	0	3,602,323
DTCA	335	510	3,001	0	31,741
Generics					
Price	2491	0.5027	0.4668	0.0378	4.5666
Share	2491	0.0053	0.0113	0.0000	0.0790
Physician ad	2491	2,012	11,220	0	99,532
DTCA	2491	0	0	0	0

Note: Observation is at the month-molecule-branded-form-firm level.

Table 1.3: Copay Estimates

Year	γ_0 Est	γ_0 S.E.	γ_1 Est	γ_1 S.E.	Plans	Total Rx
2003	−1.2393	0.0062	0.8290	0.0101	68	216920
2004	−1.2981	0.0047	0.8901	0.0086	64	129140
2005	−1.2407	0.0082	0.9064	0.0090	230	92820
2006	−1.4343	0.0141	0.9469	0.0146	295	155230
2007	−1.2624	0.0076	0.8469	0.0062	205	1431612
2008	−1.3486	0.0118	0.8682	0.0056	104	739917
2009	−1.3315	0.0182	0.8448	0.0058	204	165248

Note: Parameters are weighted OLS estimates with clustered standard errors by enrollee. Sample includes only prescriptions with 30-day supply and insurance plans with more than 10 thousand enrollees.

Table 1.4: Demand Parameters

	Est	S.E.
Nonlinear parameters:		
Nesting Parameters		
Class	0.6544	0.1995
Branded/generic	0.6323	0.2038
Molecule	0.5597	0.1255
Form	0.6650	0.0884
Proportion of high type (λ)	0.1426	0.0072
α^H	−8.6783	1.5405
α^L	−27.2243	0.9199
Linear parameters (selected):		
β^{AD}	0.1352	0.0478
β^{AC}	0.6078	0.0395
β^{ADOT}	0.3101	0.0530
β^{ACOT}	0.8623	0.2269
1m since entry	1.7625	2.2204
2m since entry	2.6343	1.9586
3m since entry	4.5392	2.2644
4m since entry	3.7120	1.9795
5m since entry	3.7589	1.9378
6m since entry	2.9701	1.7709

Note: N = 3638. Parameters are GMM estimates with heteroskedasticity-robust standard errors.

Table 1.5: Selected Cost Parameters

	Est.	S.E.
Atorvastatin (branded)	1.5261	0.1237
Rosuvastatin (branded)	1.4986	0.1243
Simvastatin (branded)	1.8890	0.1237
Simvastatin (generic)	0.2456	0.1352
Lovastatin (branded)	1.4960	0.1237
Lovastatin (generic)	0.2272	0.1457
Pravastatin (branded)	3.1775	0.1237
Pravastatin (generic)	0.5710	0.1347
One month since entry	1.3846	0.2588
One year since entry	0.3548	0.1546
Two years since entry	0.1745	0.1375

Note: N = 3638. The unit is dollar per patient-day. Adjusted R-squared = 0.71. Parameters are GMM estimates with homoskedasticity-only standard errors.

Table 1.6: Price Elasticities for June 2006

	Molecule (form)	Share	Branded											Generic		
			amlodipine/atorvastatin (tab)	ezetimibe/simvastatin (tab)	lovastatin/niacin (sa tab)	atorvastatin (tab)	fluvastatin (cap)	fluvastatin (sa tab)	lovastatin (tab)	lovastatin (sa tab)	pravastatin (tab)	rosuvastatin (tab)	simvastatin (tab)	lovastatin (tab)	pravastatin (tab)	simvastatin (tab)
Branded	amlodipine/atorvastatin (tab)	0.28%	−6.97	0.03	0.35	0.11	0.09	0.00	0.08	0.00	0.12	0.05	0.11	0.00	0.00	0.00
	ezetimibe/simvastatin (tab)	3.41%	0.20	−9.75	0.26	0.21	0.13	0.68	0.49	0.81	0.05	0.67	0.16	0.62	0.56	0.59
	lovastatin/niacin (sa tab)	0.18%	0.17	0.02	−6.97	0.04	0.04	0.06	0.03	0.08	0.05	0.02	0.04	0.01	0.01	0.01
	atorvastatin (tab)	9.48%	2.81	0.93	2.23	−5.08	2.26	0.65	2.29	0.84	3.14	1.92	3.27	0.74	0.63	0.69
	fluvastatin (cap)	0.11%	0.03	0.01	0.02	0.02	−6.90	0.07	0.02	0.01	0.03	0.01	0.03	0.01	0.00	0.00
	fluvastatin (sa tab)	0.55%	0.00	0.08	0.09	0.02	0.19	−8.49	0.05	1.24	0.00	0.07	0.01	0.08	0.07	0.08
	lovastatin (tab)	0.01%	0.00	0.00	0.00	0.00	0.00	0.00	−8.83	0.02	0.00	0.00	0.00	0.00	0.00	0.00
	lovastatin (sa tab)	0.15%	0.00	0.03	0.04	0.01	0.00	0.35	0.29	−10.77	0.00	0.02	0.00	0.08	0.02	0.03
	pravastatin (tab)	0.41%	0.15	0.01	0.12	0.15	0.12	0.00	0.10	0.01	−6.31	0.07	0.16	0.00	0.03	0.00
	rosuvastatin (tab)	1.82%	0.21	0.47	0.21	0.31	0.22	0.38	0.41	0.48	0.24	−11.00	0.29	0.44	0.38	0.41
	simvastatin (tab)	4.99%	1.58	0.37	1.25	1.76	1.26	0.25	1.20	0.32	1.75	0.95	−6.24	0.28	0.24	2.93
Generic	lovastatin (tab)	0.80%	0.00	0.05	0.01	0.01	0.01	0.06	0.08	0.20	0.00	0.05	0.01	−5.11	0.17	0.18
	pravastatin (tab)	0.39%	0.00	0.13	0.02	0.03	0.02	0.14	0.10	0.16	0.04	0.13	0.02	0.46	−27.90	0.41
	simvastatin (tab)	0.26%	0.00	0.06	0.01	0.02	0.01	0.06	0.05	0.07	0.00	0.06	0.13	0.22	0.18	−19.68

Note: Cell entries (i, j), where i indexes row and j column, give the percent change in market share of brand j with a one-percent change in price of i.

Table 1.7: Price Elasticities for Sep 2006

	Molecule (form)	Share	Branded											Generic		
			amlodipine/atorvastatin (tab)	ezetimibe/simvastatin (tab)	lovastatin/niacin (sa tab)	atorvastatin (tab)	fluvastatin (cap)	fluvastatin (sa tab)	lovastatin (tab)	lovastatin (sa tab)	pravastatin (tab)	rosuvastatin (tab)	simvastatin (tab)	lovastatin (tab)	pravastatin (tab)	simvastatin (tab)
Branded	amlodipine/atorvastatin (tab)	0.30%	−7.05	0.11	0.16	0.14	0.12	0.02	0.12	0.04	0.13	0.13	0.13	0.00	0.00	0.00
	ezetimibe/simvastatin (tab)	3.49%	0.66	−6.55	0.52	0.41	0.35	0.82	0.41	0.95	0.38	0.42	0.39	0.29	0.27	0.30
	lovastatin/niacin (sa tab)	0.17%	0.07	0.04	−5.21	0.05	0.05	0.08	0.05	0.14	0.05	0.05	0.05	0.00	0.00	0.00
	atorvastatin (tab)	8.85%	3.16	1.63	2.87	−2.33	2.96	0.47	3.12	1.02	3.46	3.57	3.67	0.02	0.02	0.02
	fluvastatin (cap)	0.09%	0.03	0.01	0.03	0.03	−5.41	0.14	0.03	0.01	0.03	0.03	0.03	0.00	0.00	0.00
	fluvastatin (sa tab)	0.47%	0.01	0.09	0.11	0.01	0.39	−7.71	0.02	1.01	0.01	0.02	0.01	0.05	0.04	0.05
	lovastatin (tab)	0.01%	0.00	0.00	0.00	0.00	0.00	0.00	−6.19	0.10	0.00	0.00	0.00	0.00	0.00	0.00
	lovastatin (sa tab)	0.13%	0.01	0.03	0.06	0.01	0.01	0.29	1.23	−8.57	0.01	0.01	0.01	0.04	0.01	0.01
	pravastatin (tab)	0.31%	0.11	0.06	0.11	0.13	0.11	0.02	0.11	0.04	−5.58	0.12	0.12	0.00	0.00	0.00
	rosuvastatin (tab)	1.90%	0.54	0.31	0.49	0.65	0.51	0.12	0.53	0.22	0.58	−5.04	0.60	0.03	0.03	0.03
	simvastatin (tab)	0.93%	0.34	0.17	0.31	0.40	0.32	0.05	0.33	0.11	0.36	0.36	−5.60	0.00	0.00	0.00
Generic	lovastatin (tab)	0.81%	0.00	0.03	0.00	0.00	0.00	0.04	0.01	0.12	0.00	0.00	0.00	−5.83	0.12	0.15
	pravastatin (tab)	0.40%	0.00	0.07	0.00	0.00	0.00	0.10	0.01	0.09	0.00	0.01	0.00	0.30	−27.49	0.31
	simvastatin (tab)	2.79%	0.00	0.32	0.01	0.01	0.01	0.49	0.04	0.45	0.00	0.04	0.00	1.70	1.37	−16.48

Note: Cell entries (i, j), where i indexes row and j column, give the percent change in market share of brand j with a one-percent change in price of i.

Table 1.8: Simulated Price, Copay, and Market Share (Counterfactual 1, No Response)

	Fraction of consumers with copay coupons				
	0%	5%	10%	30%	50%
Average price for one month supply					
Branded simvastatin	76.81	76.81	76.81	76.81	76.81
Generic simvastatin	54.43	54.43	54.43	54.43	54.42
Other branded	65.43	65.44	65.45	65.47	65.51
Other generic	31.85	31.86	31.88	31.97	32.11
Average copay for one month supply					
Copay from coupon		7.57	7.57	7.57	7.57
Branded simvastatin	17.40	17.40	17.40	17.40	17.40
Generic simvastatin	12.54	12.54	12.54	12.54	12.53
Other branded	14.94	14.94	14.94	14.95	14.96
Other generic	7.41	7.42	7.42	7.44	7.47
Market share					
Branded simvastatin	0.0103	0.0467	0.0830	0.2285	0.3740
Generic simvastatin	0.0519	0.0496	0.0472	0.0378	0.0285
Other branded	0.1584	0.1525	0.1466	0.1229	0.0993
Other generic	0.0374	0.0359	0.0344	0.0283	0.0222

Table 1.9: Change in Welfare, Profits, and Spending (Counterfactual 1, No Response)

		Fraction of consumers with copay coupons			
Change in	Original level	5%	10%	30%	50%
Consumer welfare	194,662,532	6,356,325	12,712,650	38,137,951	63,563,251
Profits: branded simvastatin	11,201,575	12,357,245	24,714,491	74,143,472	123,572,454
Profits: generic simvastatin	14,316,579	-646,221	-1,292,442	-3,877,325	-6,462,209
Profits: other branded	272,917,089	-10,194,671	-20,389,343	-61,168,028	-101,946,713
Profits: other generic	9,289,489	-376,039	-752,078	-2,256,233	-3,760,388
Insurer spending	875,612,979	129,256,873	258,513,746	775,541,237	1,292,568,728

Table 1.10: Simulated Price, Copay, and Market Share (Counterfactual 1, Price Response)

	Fraction of consumers with copay coupons				
	0%	5%	10%	30%	50%
Average price for one month supply					
Branded simvastatin	76.81	96.01	96.01	95.99	96.01
Generic simvastatin	54.43	54.43	54.43	54.43	54.42
Other branded	65.43	67.18	67.58	69.40	71.62
Other generic	31.85	31.81	31.82	31.81	31.88
Average copay for one month supply					
Copay from coupon		5.76	5.76	5.76	5.76
Branded simvastatin	17.40	21.50	21.50	21.50	21.50
Generic simvastatin	12.54	12.54	12.54	12.54	12.53
Other branded	14.94	15.32	15.40	15.80	16.28
Other generic	7.41	7.40	7.41	7.40	7.42
Market share					
Branded simvastatin	0.0103	0.0491	0.0952	0.2800	0.4653
Generic simvastatin	0.0519	0.0496	0.0472	0.0373	0.0272
Other branded	0.1584	0.1547	0.1462	0.1131	0.0827
Other generic	0.0374	0.0358	0.0341	0.0272	0.0201

Table 1.11: Change in Welfare, Profits, and Spending (Counterfactual 1, Price Response)

		Fraction of consumers with copay coupons			
Change in	Original level	5%	10%	30%	50%
Consumer welfare	194,662,532	6,784,959	18,931,200	67,617,547	116,405,789
Profits: branded simvastatin	11,201,575	58,045,503	119,951,619	367,570,401	616,423,199
Profits: generic simvastatin	14,316,579	-606,152	-1,275,920	-3,990,257	-6,797,782
Profits: other branded	272,917,089	10,681,565	-1,328,374	-49,823,252	-98,777,783
Profits: other generic	9,289,489	-382,607	-806,373	-2,523,202	-4,300,805
Insurer spending	875,612,979	219,759,515	436,442,418	1,305,670,449	2,183,617,216

Table 1.12: Simulated Price, Copay, and Market Share (Counterfactual 2, No Response)

	Fraction of consumers with copay coupons				
	0%	5%	10%	30%	50%
Average price for one month supply					
Branded simvastatin	76.81	76.81	76.81	76.81	76.81
Generic simvastatin	54.43	54.43	54.43	54.43	54.41
Other branded	65.43	65.60	65.76	66.44	67.17
Other generic	31.85	31.86	31.88	31.99	32.18
Average copay for one month supply					
Copay from coupon		7.45	7.45	7.45	7.45
Branded simvastatin	17.40	17.40	17.40	17.40	17.40
Generic simvastatin	12.54	12.54	12.54	12.54	12.53
Other branded	14.94	14.97	15.01	15.16	15.32
Other generic	7.41	7.42	7.42	7.44	7.49
Market share					
Branded simvastatin	0.0103	0.0481	0.0858	0.2369	0.3880
Generic simvastatin	0.0519	0.0491	0.0464	0.0353	0.0243
Other branded	0.1584	0.1571	0.1558	0.1506	0.1454
Other generic	0.0374	0.0356	0.0338	0.0266	0.0194

Table 1.13: Change in Welfare, Profits, and Spending (Counterfactual 2, No Response)

		Fraction of consumers with copay coupons			
Change in	Original level	5%	10%	30%	50%
Consumer welfare	194,662,532	6,683,085	13,366,170	40,098,510	66,830,850
Profits: branded simvastatin	11,201,575	12,895,452	25,790,905	77,372,715	128,954,524
Profits: generic simvastatin	14,316,579	-761,178	-1,522,355	-4,567,066	-7,611,776
Profits: other branded	272,917,089	-748,909	-1,497,818	-4,493,453	-7,489,089
Profits: other generic	9,289,489	-447,007	-894,014	-2,682,042	-4,470,070
Insurer spending	875,612,979	152,541,838	305,083,676	915,251,026	1,525,418,377

Table 1.14: Simulated Price, Copay, and Market Share (Counterfactual 2, Price Response)

	Fraction of consumers with copay coupons				
	0%	5%	10%	30%	50%
Average price for one month supply					
Branded simvastatin	76.81	95.98	95.98	96.00	96.00
Generic simvastatin	54.43	54.43	54.43	54.42	54.42
Other branded	65.43	67.41	68.06	70.63	72.81
Other generic	31.85	31.82	31.82	31.84	31.88
Average copay for one month supply					
Copay from coupon		5.72	5.72	5.72	5.72
Branded simvastatin	17.40	21.49	21.49	21.50	21.50
Generic simvastatin	12.54	12.54	12.54	12.54	12.54
Other branded	14.94	15.36	15.50	16.06	16.54
Other generic	7.41	7.41	7.41	7.41	7.42
Market share					
Branded simvastatin	0.0103	0.0499	0.0968	0.2847	0.4726
Generic simvastatin	0.0519	0.0492	0.0464	0.0347	0.0228
Other branded	0.1584	0.1605	0.1576	0.1483	0.1425
Other generic	0.0374	0.0356	0.0336	0.0254	0.0170

Table 1.15: Change in Welfare, Profits, and Spending (Counterfactual 2, Price Response)

		Fraction of consumers with copay coupons			
Change in	Original level	5%	10%	30%	50%
Consumer welfare	194,662,532	6,659,462	18,609,664	66,841,642	116,015,374
Profits: branded simvastatin	11,201,575	59,100,286	122,071,821	374,466,147	626,872,638
Profits: generic simvastatin	14,316,579	-727,714	-1,514,159	-4,720,815	-8,022,187
Profits: other branded	272,917,089	23,454,134	24,671,679	30,853,786	38,207,213
Profits: other generic	9,289,489	-460,303	-958,067	-2,988,045	-5,078,644
Insurer spending	875,612,979	246,868,593	490,774,719	1,472,669,756	2,462,710,258

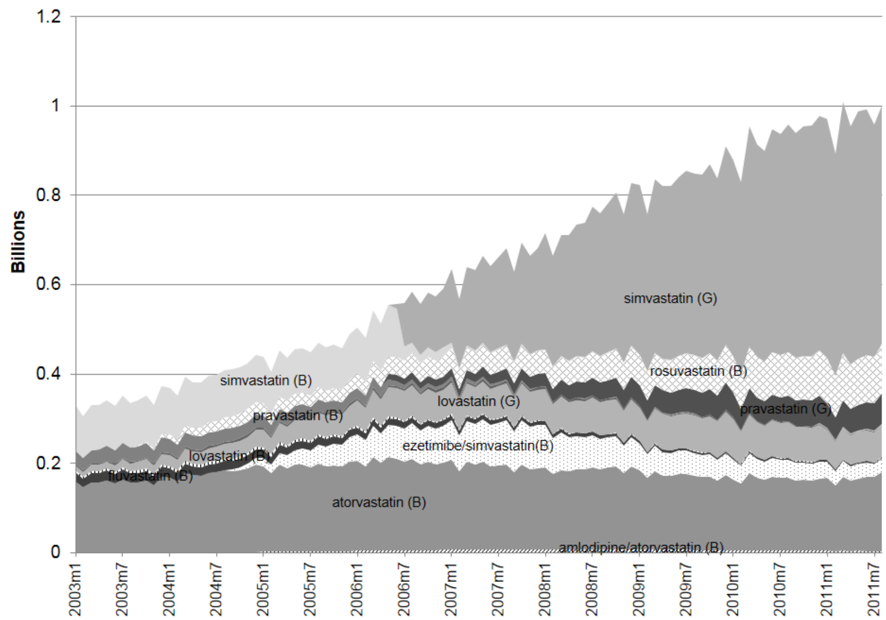


Figure 1.1: Sales (Patient-Days)

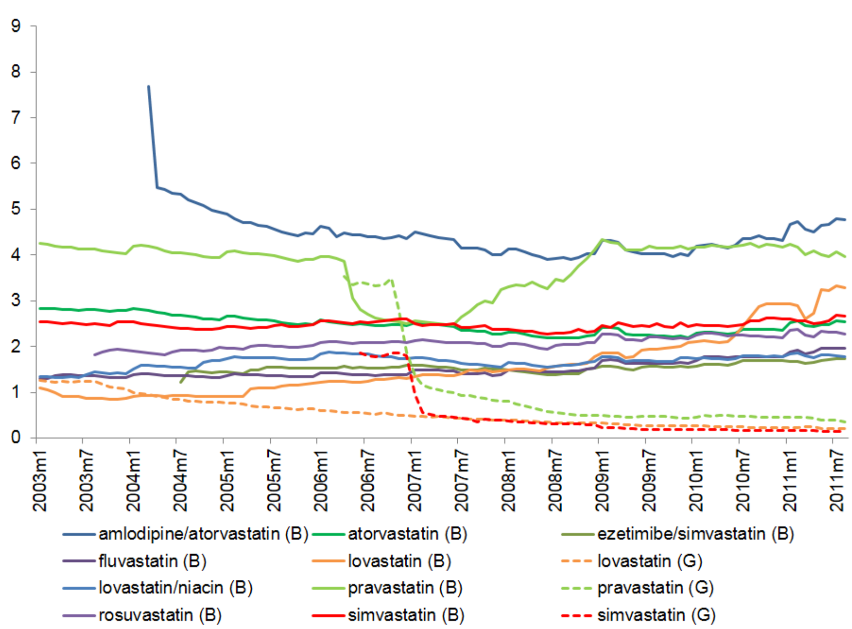


Figure 1.2: Price per Patient-Day

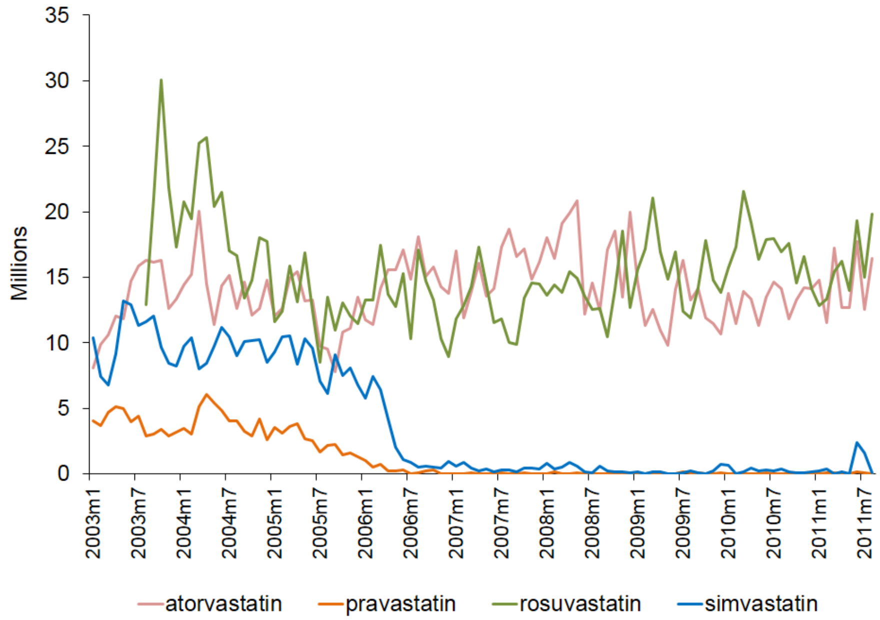


Figure 1.3: Spending on Advertising to Physicians

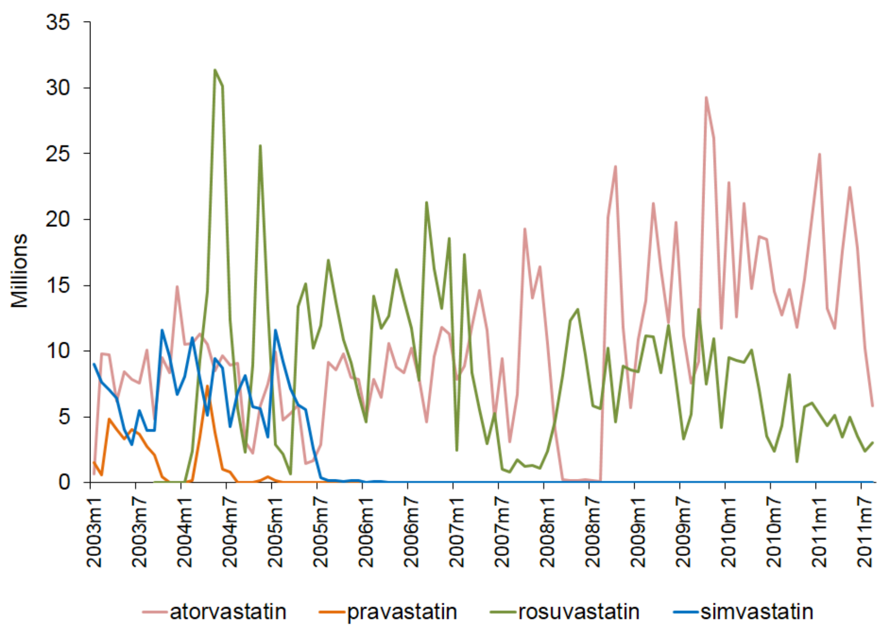


Figure 1.4: Spending on Direct-to-Consumer Advertising

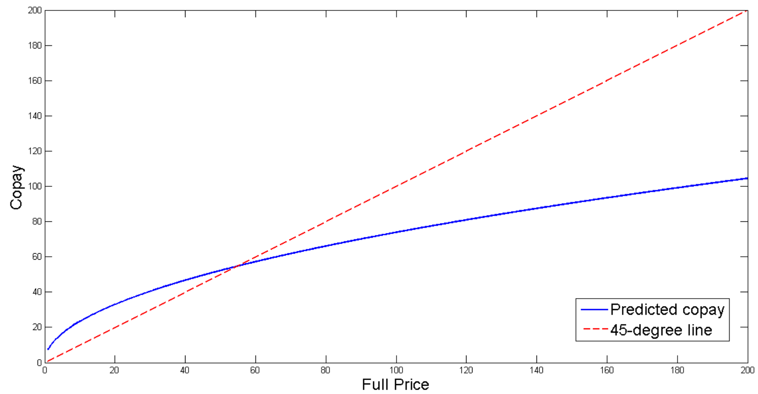


FIGURE 1.5: Example: Relationship between Full Price and Copay

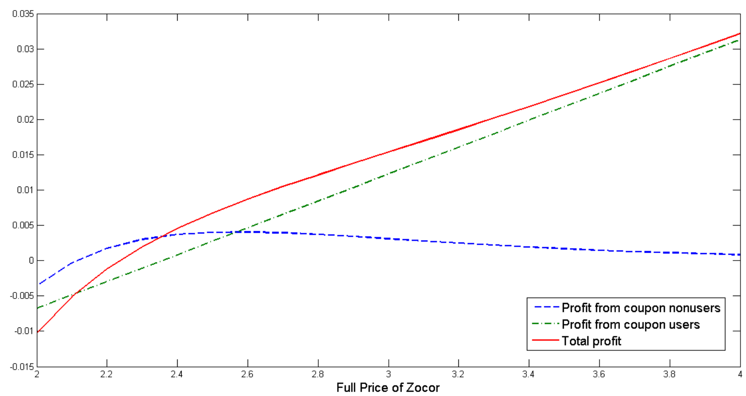


FIGURE 1.6: Example: Profit of Merck vs. Zocor's Full Price

2

Brand Loyalty and Learning in Pharmaceutical Demand

2.1 Introduction

In recent years, branded drug manufacturers are facing generic competition with many of their blockbuster drugs. According to IMS Health, the patents of six of the ten best-selling prescription drugs in the US expired in 2011 and 2012. To retain revenue after patent expiration, branded drug manufacturers have employed several strategies, including pay-for-delay agreements (paying generic companies not to bring lower-cost alternatives to market) and copay coupons (distributing a card directly to patients to lower their out-of-pocket costs). The success of these strategies depends not only on how many patients drug manufacturers can attract today, but also on how many of the patients will stick to the brands tomorrow. Thus, patient stickiness plays an important role in these marketing strategies and understanding the sources of patient stickiness in pharmaceuticals can help manage the programs.

This paper estimates the demand for pharmaceuticals by incorporating brand loyalty and learning to study patient stickiness. Using the micro-level databases for cholesterol-lowering drugs, I look at the switch patterns of patients with different lengths of treatment. The data shows inter-molecule switching probabilities are higher in early prescriptions and quickly decrease in a treatment, suggesting the existence of switching costs and learning about molecules. Switching costs make a patient loyal to a drug and lower the probability of choosing a different drug next time.[1] On the other hand, patients learn the effects of drugs by taking it repeatedly and will probably stop experimenting once they find a drug that works well.

I disentangle the two effects in a multinomial logit model and estimate the effects using patients' prescription history. The estimates suggest high switching costs and strong learning effects at the molecule level in the cholesterol lowering drug market. Inter-molecule switching costs are larger than intra-molecule switching costs. The effect of learning about molecules gets larger as a patient takes a drug longer. In addition, switching costs largely raise the probability of choosing the same drug for new patients. Both learning and switching costs contribute to experienced patients' stickiness in the long run.

The paper adds to the literature on consumer dynamics in pharmaceuticals by exploiting the rich micro-level data and identifying effects of brand loyalty and learning. Coscelli (2000) uses a prescription level dataset from the Italian markets to study the relationship between probabilities of switching brands and patient and doctor attributes. He shows habit persistence of doctors and patients in prescription choices. Crawford and Shum (2005) use the same dataset to study inter-molecule choice by considering consumer learning and patient heterogeneity. However, the lack of price variation in Italian markets makes it impossible to identify the switching cost in their structural analysis. My datasets from the US pharmaceutical markets can overcome this problem since drug manufacturers are allowed to compete in price in

[1] In the paper, I use "switching cost" and "loyalty" interchangeably. Klemperer (1995) argues that brand loyalty can create switching costs. Studying the actual sources of switching costs in pharmaceuticals is beyond the scope of this paper.

US. Moreover, my datasets have information on insurance plan design and this reveals variation in out-of-pocket costs for drugs with different status in an insurance plan. The price faced by a patient is her cost share rather than the full price. With the information on out-of-pocket costs, I can more adequately estimate the price elasticities.

Dalen et al. (2011) and Lundin (2000) estimate binary choice models based on prescription level data with price information. They find price difference between branded drugs and generics an important factor in generic substitution. Both have limitations on the use of patients' prescription history. Dalen et al. (2011) ignores past use and Lundin (2000) only considers switching cost by including the last prescription for each observation in the demand model. My paper takes into account both switching costs, which incur when patient chooses a drug different than the choice last time, and learning, which is revealed in the choice of the same drugs multiple times. Furthermore, I consider learning at the molecule level and the version (branded/generic) level. Incorporating learning about the generic version of a molecule in the demand model can help understand intra-molecule switching after patent expiration.

Finally, the paper also contributes to the empirical study of consumer brand loyalty in the marketing literature. Past studies in this field (Krishnamurthi and Raj, 1991; Erdem, 1996; Allenby and Lenk, 1995; Keane, 1997; Dubé et al., 2010; Osborne, 2011; Bronnenberg et al., 2012) focus on brand loyalty and state dependence in consumer goods. My paper looks at pharmaceuticals, which have more than 300 billion US dollar sales in a year. Consumer dynamics in this industry can be of interest to managers and policy makers. For example, drug manufacturers may want to know how large switching costs are for patients to try a drug when they are marketing a new product. Policy makers may be interested in the learning process in a treatment, which can provide directions to improve patient compliance or medication adherence.

The rest of the paper is organized as follows. Section 2.2 describes data and presents relevant statistics and trends. Section 2.3 develops models for demand. Section 2.4 discusses estimation strategies and results. Section 2.5 concludes.

2.2 Data

The data are obtained from the MarketScan Commercial Claims and Encounters Database and the MarketScan Benefit Plan Design Database through National Bureau of Economic Research (NBER). The MarketScan Databases are constructed from privately insured paid medical and prescription drug claims. There are about 100 payers and more than 500 million claim records in the Databases. The Commercial Claims and Encounters Database provides the prescription-level data on date of service, drug characteristics, days of supply, full price, out-of-pocket cost, patient age, and patient gender. The "plan key" variable in the data can be used to link the Commercial Claims and Encounters Database to the Benefit Plan Design Database, which provides the health insurance information for the patient, including copayment and coinsurance for drugs in different tiers.

My work focuses on the market of cholesterol-lowering drugs (statins) from 2005 to 2006. There are several reasons to analyze the market for the years. First, cholesterol-lowering drug is the largest therapeutic class in the US by spending in 2006 and had approximately 22 billion US dollar sales and 210 million prescriptions in the year.[2] Cholesterol-lowering drugs can help lower rates of low-density lipoprotein (LDL) cholesterol in the blood and are usually used repeatedly, which gives the opportunity to observe the choices a patient made for several periods. This enables me to examine the dynamic aspect of pharmaceutical demand. Second, one of the best-selling drugs in this class, Zocor, lost its patent protection in mid 2006. With

[2] *The Use of Medicines in the United States: Review of 2010*, IMS Institute.

generic entry after patent expiration, I am able to see how patients respond to the availability of low-cost alternatives. Third, the drug manufacturer Pfizer employed aggressive strategies in this market, including pay-for-delay agreements and copay coupons, to retain revenue from Lipitor after its patent expiration in 2011. Thus, understanding consumer switch between branded drugs and generics in this market can shed light on the possible outcomes of Pfizer's strategies.

I take three major steps to prepare the data to be used in this paper. First, I keep individuals who have at least one prescription recorded in both 2005 and 2006 to make sure the individuals in the sample are enrolled in both years. Second, I do not observe when a patient started to take drugs. To avoid difficulty associated with the left-censoring, I include only patients with their first prescriptions observed after June 30, 2005. By doing so, I assume that patients who haven't took drugs for more than half a year are new in the market.[3] Third, since I only observe the out-of-pocket cost for the drug purchased, I calculate the out-of-pocket costs for all available drugs faced by a patient by taking the average out-of-pocket cost for each drug across prescriptions belonging to the patient's insurance plan. For some small insurance plans, there are no prescriptions for certain drugs and thus the average out-of-pocket costs for those drugs in the plan are not available. I drop patients in the plans with missing out-of-pocket costs. The filters together left me with 18,316 patients and 121,033 prescriptions.

Table 2.1 presents the summary statistics. There are six molecules with a market share greater than 3% and two of them have generic equivalents. Generic simvastatin entered in mid 2006 and it is the only new choice in the sample. The average full price for a 30-day supply for branded drugs fall between \$78 and \$118, while the

[3] In the sample, prescriptions with more than six months of supply are rare. Thus, based on the definition of a new patient, the cleaned data mostly covers individuals who entered this market for the first time ever and those who had not used drugs in this therapeutic class for a long time. The latter type of patients is assumed not to retain information on anything which happened before.

average full price for generic lovastatin is only \$21.4. The price for generic simvastatin is quite high since there was limited competition among generic manufacturers during the first few months of Zocor's patent expiration.[4] The out-of-pocket costs are the net payments from patients and they are about 22% to 30% the full price of branded drugs. For the generics, the out-of-pocket costs are much lower since insurance companies usually ask for a lower cost share to induce patients to choose less expensive generic alternatives.

Next, I discuss evidence of consumer dynamics in the data. I define days of treatment as the minimum of the days of supply for a prescription and the number of days between the prescription and the next one. Also, following Crawford and Shum (2005), I define a drug "spell" to be a sequence of one or more prescriptions to a single molecule or molecule-version. Table 2.2 shows that an average patient has 6.7 prescriptions and 10.7 months of treatment in the one year and a half window. The average number of spells for molecules is about 1.2 and for molecule-versions is about 1.3, implying that patients in the sample do not change drugs very frequently. The patterns of inter-molecule switching shown in Table 2.3 and Table 2.4 are consistent with this finding. The switch probabilities are high at the beginning of a treatment. Less than 3% of the patients switch to another molecule after the fifth prescription or 120 days of treatment. The high switch probabilities are probably resulted from experimentation. At the beginning of a treatment, doctors are trying different drugs to find the best fit for their patients. Once they learn that a drug matches a patient well, they would keep prescribing the drug for the patient.

Finally, I look at patients' switching to generic simvastatin after Zocor's patent expiration on June 23, 2006. Figure 2.1 shows that over 55% of patients who took

[4] US Food and Drug Administration (FDA) granted a 180-day exclusivity to Teva and Ranbaxy to sell generic simvastatins since they are the first challengers of Zocor's primary US patent. Thus, during the 180 days after Zocor's patent expiration, there were only three generic manufacturers: the two independent generic producers and Dr. Reddy's Laboratories, which received a license from Merck to sell "authorized" generic simvastatin.

Zocor before and still want to take Zocor choose generic simvastatin over the branded version when generic simvastatins are available. This seems to imply that the low out-of-pocket costs for generic simvastatin make them very attractive. Figure 2.2, on the other hand, shows that those who have taken Zocor for more than 10 months are less willing to switch to the generic version. About half of them stay with branded Zocor, compared to less than 35% of those with less experience with Zocor. Those who have used Zocor for more than 10 months may face lower out-of-pocket costs for branded Zocor and/or have learned that the branded version is something they really like after using it for a long time. The result suggests that learning at the version level may take longer. To disentangle the price effect from the learning effect, we need a model that can handle both at the same time.

2.3 Model

Consider a patient i who goes to a doctor in period t to seek treatment for high cholesterol. After diagnosing the patient and observing her copayments for each drug (p^c_{ijt}) as well as her prescription history, the doctor selects drug j from the choice set J_t for the patient.[5] A drug here is a combination of a molecule and a version (branded or generic), e.g. branded atorvastatin, generic simvastatin, etc. The utility patient i derives from drug j in period t is

$$\begin{aligned} U_{ijt} &= \zeta_j + \alpha p^c_{ijt} + \beta_1 I\left(m_{ijt-1} = 1\right) + \beta_2 I\left(d_{ijt-1} = 1\right) \\ &\quad + X^m_{ijt}\beta_3 + X^d_{ijt}\beta_4 + \epsilon_{ijt} \qquad (2.1) \\ &= u_{ijt} + \epsilon_{ijt}, \end{aligned}$$

[5] Due to the data limitations, I am not able to distinguish the roles of doctors and patients. Thus, I assume that there is no agency issue and that doctors always act in the best interest of the patients.

where ζ_j is the base utility from drug j (drug fixed effect),[6] p^c_{ijt} the out-of-pocket cost for drug j paid by patient i in period t, m_{ijt} a dummy variable equal to one if the doctor chooses the molecule of drug j for patient i in period t, d_{ijt} a dummy variable equal to one if the doctor chooses drug j for patient i in period t, X^m_{ijt} a vector of dummy variables for past use of drug j's molecule, X^d_{ijt} a vector of dummy variables for past use of drug j, and ϵ_{ijt} an idiosyncratic error term assumed to follow the Type I Extreme Value distribution.

The variable $I(m_{ijt-1} = 1)$ is an indicator which equals one if the doctor chose the same molecule of drug j for the patient in the last period, and zero otherwise. The coefficient β_1 thus accounts for the dynamic behaviors of switching costs or experimenting. A positive β_1 induces a inter-molecule switching cost while a negative β_1 implies that doctors like to experiment with various molecules to find the best fit for the patient. That is, if $\beta_1 > 0$, patient i's utility is greater if she takes the same molecule in a row. If $\beta_1 < 0$, the patient prefers to try something different than her last drug choice.

Similarly, $I(d_{ijt-1} = 1)$ is an indicator which equals one if the doctor chose the same drug (molecule-version) for the patient in the last period, and zero otherwise. If $I(m_{ijt-1} = 1) = 1$ but $I(d_{ijt-1} = 1) \neq 1$, the doctor chooses the same molecule this period as the molecule for the last period, but she chooses a different version of the molecule. This will happen when a doctor selected a branded drug last time but chooses its generic equivalent this time. If $I(m_{ijt-1} = 1) = 1$ and $I(d_{ijt-1} = 1) = 1$, the doctor chooses the same molecule and the same version for patient i in a row. If $\beta_1 + \beta_2 > 0$, the patient has inter- and intra-molecule switching cost and she has greater utility choosing the same molecule and version successively. If $\beta_1 + \beta_2 > 0$ and $\beta_2 < 0$, the patient prefers to use the same molecule but try a different version.

[6] Since only differences in utility matters in this discrete choice model and we can only identify relative fixed effects, I normalize the fixed effect for the "other" drugs to be zero.

X_{ijt}^{m} captures the effect of learning about the molecule of a drug. By including dummy variables for different lengths of using a molecule, measured by number of prescriptions or treatment days, we will be able to see whether a patient that receives multiple prescriptions for the same molecule or several days of treatment with the molecule has learned that the molecule is effective in lowering her cholesterol level. For example, a positive coefficient on the dummy $I\left(2 \leqslant \sum_{k=1}^{t-1} m_{ijk} < 4\right)$ implies that the patient has tried the drug two or three times and really likes it. In addition, comparing the coefficients on $I\left(2 \leqslant \sum_{k=1}^{t-1} m_{ijk} < 4\right)$ and $I\left(4 \leqslant \sum_{k=1}^{t-1} m_{ijk} < 6\right)$ will reveal the difference in the degree of learning between a patient with 2 or 3 past prescriptions and a patient with 4 or 5 past prescriptions for the molecule.

On the other hand, X_{ijt}^{d} captures the effect of learning about the version of a drug. X_{ijt}^{d} includes dummy variables for different lengths of using a drug (molecule-version). If the coefficient on $I\left(2 \leqslant \sum_{k=1}^{t-1} m_{ijk} < 4\right)$ is positive and the coefficient on $I\left(2 \leqslant \sum_{k=1}^{t-1} d_{ijk} < 4\right)$ is positive, a patient with 2 or 3 prescriptions for the drug likes the molecule as well as the version. However, If the coefficient on $I\left(2 \leqslant \sum_{k=1}^{t-1} m_{ijk} < 4\right)$ is positive and the coefficient on $I\left(2 \leqslant \sum_{k=1}^{t-1} d_{ijk} < 4\right)$ is negative, the patient with 2 or 3 prescriptions for the drug likes the molecule but prefers a different version. Moreover, we can study how lengths of treatment affect preference for a molecule and version by comparing the coefficients on the dummies for different lengths of treatment with a drug.

The assumption that ϵ_{ijt} are independent and identically (i.i.d.) distributed Type 1 Extreme Value gives the likelihood of patient i choosing drug j in period t

$$\Pr\left(d_{ijt} = 1 \mid \mathbf{X}_{it}\right) = \frac{\exp(u_{ijt})}{\sum_{l=1}^{J_t} \exp(u_{ilt})} \quad j = 1, 2, ..., J_t, \tag{2.2}$$

where $\mathbf{X}_{it}$ denotes the set of conditioning variables for individual i in period t.

2.4 Estimation

With the specification of the likelihood of d_{ijt} given $\mathbf{X}_{it}$, we can estimate the parameters using maximum likelihood (MLE). The log-likelihood can be written as

$$\log\left(L(\theta)\right) = \sum_{t=1}^{T}\sum_{i=1}^{N_t}\sum_{j=1}^{J_t} I\left(d_{ijt}=1\right)\log\left(\frac{\exp(u_{ijt})}{\sum_{l=1}^{J_t}\exp(u_{ilt})}\right), \tag{2.3}$$

where θ is the vector of the unknown parameters. McFadden (1973) has shown that the log-likelihood function is globally concave and discussed the conditions for MLE to be consistent and asymptotically normal in this application.

The out-of-pocket cost variable (p^c_{ijt}) is endogenous since drug with a better (unobserved) quality can give the patient greater utility and cost more. To alleviate the bias associated with this endogeneity problem, I make use of the panel structure of the data by including the drug fixed effect in estimation. Drug quality is not observable but is assumed to be fixed over time. Thus, adding drug fixed effect can help solve the problem from endogenous out-of-pocket costs. Identification of the other parameters relies on the variation in the corresponding variables.

Table 2.5 summarizes the estimation results. First note that the price coefficient is more negative in the models with drug fixed effect, which implies that the coefficient is overestimated if we do not add the drug fixed effect. Based on the estimates from Model (2), the average own price elasticities are between -0.30 and -0.08, and average cross price elasticities are between 0.006 and 0.4. Additionally, the estimates for the coefficients $I(d_{ijt-1}=1)$ and $I(m_{ijt-1}=1)$ are both positive and significant. This means that there is a positive cost if a patient switch to a different molecule or version. Other things being equal, patients prefer to take the same drug in this period as the drug they took in the last period. Furthermore, inter-molecule switching cost is higher than intra-molecule switching cost since the sum of the two estimated coefficients is larger than the estimated coefficient on $I(m_{ijt-1} = 1)$.

Results from Model (2) and (4) show that the effects of learning about molecules are all positive and significant. Also, the longer a patient takes a molecule, the more likely the patient will keep choosing the molecule. In contrast, the effect of learning about the version of a drug is not significant for patients who take the drug for less than 10 prescriptions or 300 days. Only those who have more than 10 past prescriptions or 300 days with a drug demonstrate a significant effect of learning about versions. The effect of learning about versions is weaker than the effect of learning about molecules. In other words, patients in general do not learn to distinguish between versions of a molecule unless they have taken the drug for a very long time. This is consistent with our findings in the data analysis. When generics enter the market, the patients with less experience with Zocor are more likely to switch to generic simvastatin.

The effect of switching costs and learning can be further illustrated in changes in the probability of choosing a drug for patients with different lengths of taking the drug. Table 2.6 presents the predicted probabilities of choosing a drug in successive prescriptions with different lengths of treatment with the drug. The probabilities are calculated using the estimates from Model (2). The first column in the table shows that the choice probabilities vary a lot for the first prescription. New patients' choice for the first prescription is based on the out-of-pocket costs and the mean utility for each drug. For the second prescription, the probability of choosing the same drug increases by more than 40 percentage points. The probability of choosing blockbuster drug Lipitor jumps from 23.8% to 96.6% after the first use. The probability of choosing the less attractive drug Pravachol also largely increases from 3.5% to more than 76.8%. High switching costs thus make a patient quite likely to stay with a drug even when they use the drug just once.

The probabilities of choosing the same drug for experienced patients are resulted from both learning and switching costs. In Table 2.6, the predicted probabilities in general increase with the lengths of treatment. Also, the difference in the choice probabilities among drugs becomes smaller as a patient gets more experience with a drug. Those who have chosen most of the drugs for more than ten successive prescriptions will be very unlikely to change their mind for the next prescription. Learning about molecules and switching costs together lead to high probabilities of choosing a drug in a row.

It is interesting to note the spillover learning effect with Zocor. For a patient who has taken Zocor for two or three prescriptions, the probability of choosing Zocor again is less than 50%, compared to more than 80% for the other drugs. This is because patients' learning about Zocor's molecule (simvastatin) benefits the generic competitors. As discussed above, the effect of learning about molecules is stronger than the effect of learning about versions. Also, intra-molecule switching cost is lower. The low cost generic simvastatins is thus very attractive to the experienced Zocor users and lowers the probability they stick with Zocor.

To sum up, the estimates suggest high switching costs and strong learning effect at the molecule level in the cholesterol lowering drug market. Effects of learning about molecules get larger as a patient takes a drug longer. Effects of learning about drug versions are significant only for the patients who have used a drug for a long time. In addition, switching costs raise the probability of choosing the same drug for new patients. Both learning and switching costs contribute to experienced patients' stickiness.

2.5 Conclusion

The paper estimates a multinomial logit demand model with brand loyalty and learning using data from cholesterol lowering drug markets in US. The estimation results imply that there are high switching costs and strong learning effect at the molecule level. The longer a patient takes a drug, the stronger the learning effect is. In contrast, effects of learning about drug versions are significant only for patients who have used a drug over 10 months. Also, the predicted choice probabilities show that switching costs increase probabilities of choosing the same drug for the patients who just start a treatment. Experienced patients stick to a drug because they have learned the molecule works well and there are high switching costs.

These results have several implications for drug manufacturers, insurance companies, and health care policy makers. First, when rolling out a new product in a market with existing competitors, drug manufacturers should lower the high inter-molecule switching costs to encourage patients to try it. They can provide free samples to reduce the financial costs or invest in advertising to lower the information costs. Second, to encourage generic use and create savings for insurance plans, insurance companies can consider providing more incentive to those who have never tried the molecules of the generics. Since it is easier for experienced users of a molecule to switch to generics after patent expiration, insurance companies can focus more on new users by further lowering or eliminating their copayment for their first prescription for the molecule. Finally, policy makers can provide subsidy to new patients for their first few prescriptions to improve medication adherence. Once patients learn the effectiveness of a drug, they can better comply with the treatment.

Table 2.1: Summary Statistics

Product	Molecule	Branded/ generic	Date of entry	In-sample market share	Ave full price for 30 days	Ave copay for 30 days
Lipitor	atorvastatin	B	Jan 97	44.34	84.9	18.6
Vytorin	ezetimibe/simvastatin	B	Jul 04	17.83	78.6	18.4
Crestor	rosuvastatin	B	Aug 03	10.60	79.7	19.7
Zocor	simvastatin	B	Jan 92	9.78	115.3	30.0
Lovastatin	lovastatin	G	Feb 02	8.23	21.4	6.3
Pravachol	pravastatin	B	Nov 91	3.56	117.7	21.5
Simvastatin	simvastatin	G	Jun 06	3.00	93.6	9.1
5 others	-	B	$<$ Jul 02	2.67	80.7	21.3

Note: The level of observations is prescription.

Table 2.2: Overall Statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Number of Rx	18,316	6.69	3.90	1	22
Days of treatment	18,316	321.43	136.94	8	838
Number of spells (molecule)	18,316	1.19	0.49	1	13
Number of spells (molecule-version)	18,316	1.25	0.54	1	13
Age	18,316	52.27	8.20	0	64
Female	18,316	0.46	0.50	0	1

Note: The level of observations is patient.

Table 2.3: Inter-molecule Switch by Prescriptions

Rx No.	# patients	# patients with a molecule switch	Percentage
2	17737	1045	5.9%
3	16377	710	4.3%
4	14400	507	3.5%
5	12073	363	3.0%
6	9578	247	2.6%
7	7194	171	2.4%
8	5653	91	1.6%
9	4634	67	1.4%
10	3864	66	1.7%
>10	11769	128	1.1%

Table 2.4: Inter-molecule Switch by Days of Treatment

Days of Treatment	# patients	# patients with a molecule switch	Percentage
30 - 59	18290	741	4.1%
60 - 89	17665	476	2.7%
90 - 119	16952	626	3.7%
120 - 149	16040	274	1.7%
150 - 179	15193	236	1.6%
180 - 209	13836	263	1.9%
210 - 239	12938	138	1.1%
240 - 269	11825	145	1.2%
270 - 299	10437	132	1.3%
>= 300	10152	262	2.6%

Table 2.5: Multinomial Logit Estimates of Drug Choices

Variable	(1) Est.	(1) S.E.	(2) Est.	(2) S.E.	(3) Est.	(3) S.E.	(4) Est.	(4) S.E.
p^c	-0.0202	0.0007	-0.0316	0.0011	-0.0191	0.0007	-0.0312	0.0011
$I(d_{ijt-1}=1)$	1.7061	0.0572	1.6703	0.0632	1.8782	0.0710	1.6540	0.0701
$I(m_{ijt-1}=1)$	2.8287	0.0607	2.8342	0.0670	2.6712	0.0747	2.8335	0.0743
Lengths of treatment in Rx								
$I\left(2 \leqslant \sum_{k=1}^{t-1} m_{ijk} < 4\right)$	1.5500	0.0890	0.6318	0.1008				
$I\left(4 \leqslant \sum_{k=1}^{t-1} m_{ijk} < 6\right)$	1.9101	0.0966	0.8306	0.1097				
$I\left(6 \leqslant \sum_{k=1}^{t-1} m_{ijk} < 8\right)$	2.5066	0.1052	1.4589	0.1275				
$I\left(8 \leqslant \sum_{k=1}^{t-1} m_{ijk} < 10\right)$	2.7167	0.1357	1.5082	0.1570				
$I\left(10 \leqslant \sum_{k=1}^{t-1} m_{ijk}\right)$	2.8396	0.1393	1.5059	0.1511				
$I\left(2 \leqslant \sum_{k=1}^{t-1} d_{ijk} < 4\right)$	-0.9639	0.0831	-0.1226	0.0964				
$I\left(4 \leqslant \sum_{k=1}^{t-1} d_{ijk} < 6\right)$	-0.8937	0.0853	0.1341	0.1019				
$I\left(6 \leqslant \sum_{k=1}^{t-1} d_{ijk} < 8\right)$	-1.1586	0.0828	-0.2033	0.1128				
$I\left(8 \leqslant \sum_{k=1}^{t-1} d_{ijk} < 10\right)$	-1.0550	0.0982	0.1095	0.1296				
$I\left(10 \leqslant \sum_{k=1}^{t-1} d_{ijk}\right)$	-0.8264	0.1053	0.5364	0.1268				
Lengths of treatment in days								
$I\left(60 \leqslant \sum_{k=1}^{t-1} D_{ik} \times m_{ijk} < 120\right)$					1.1241	0.1193	0.2119	0.1250
$I\left(120 \leqslant \sum_{k=1}^{t-1} D_{ik} \times m_{ijk} < 180\right)$					2.2770	0.1203	1.0738	0.1317
$I\left(180 \leqslant \sum_{k=1}^{t-1} D_{ik} \times m_{ijk} < 240\right)$					2.6854	0.1106	1.3961	0.1223
$I\left(240 \leqslant \sum_{k=1}^{t-1} D_{ik} \times m_{ijk} < 300\right)$					2.7755	0.1213	1.4988	0.1304
$I\left(300 \leqslant \sum_{k=1}^{t-1} D_{ik} \times m_{ijk}\right)$					2.4194	0.1182	1.1382	0.1266
$I\left(60 \leqslant \sum_{k=1}^{t-1} D_{ik} \times d_{ijk} < 120\right)$					-0.6687	0.1139	0.1957	0.1204
$I\left(120 \leqslant \sum_{k=1}^{t-1} D_{ik} \times d_{ijk} < 180\right)$					-1.3257	0.1110	-0.1596	0.1243
$I\left(180 \leqslant \sum_{k=1}^{t-1} D_{ik} \times d_{ijk} < 240\right)$					-1.6931	0.0978	-0.3776	0.1144
$I\left(240 \leqslant \sum_{k=1}^{t-1} D_{ik} \times d_{ijk} < 300\right)$					-1.4810	0.1002	-0.1522	0.1167
$I\left(300 \leqslant \sum_{k=1}^{t-1} D_{ik} \times d_{ijk}\right)$					-0.8956	0.0980	0.4285	0.1152
Drug fixed effect	No		Yes		No		Yes	
# obs	121033		121033		121033		121033	
Log-likelihood value	-56940.9		-50349.5		-56993.4		-50422.8	

Note: Homoskedasticity-robust standard errors are reported. D_{ik} is days of treatment for patient i in period k.

Table 2.6: Predicted Probabilities of Choosing the Same Drug

	Cumulative Rx with Product						
Product	0	1	2 - 3	4 - 5	6 - 7	8 - 9	>= 10
Crestor	0.0800	0.8872	0.9290	0.9538	0.9650	0.9754	0.9838
Lipitor	0.2381	0.9658	0.9792	0.9867	0.9900	0.9930	0.9954
Lovastatin	0.0538	0.8370	0.8953	0.9309	0.9475	0.9628	0.9754
Other	0.0265	0.7109	0.8036	0.8658	0.8962	0.9253	0.9499
Pravachol	0.0353	0.7681	0.8465	0.8968	0.9208	0.9435	0.9623
Simvastatin	0.3707	0.9489	0.9505	0.9625	0.9520	0.9646	0.9766
Vytorin	0.1222	0.9264	0.9544	0.9706	0.9779	0.9845	0.9898
Zocor	0.0735	0.4919	0.4709	0.5370	0.4572	0.5355	0.6385

Note: Choice probabilities are evaluated at mean out-of-pocket costs.

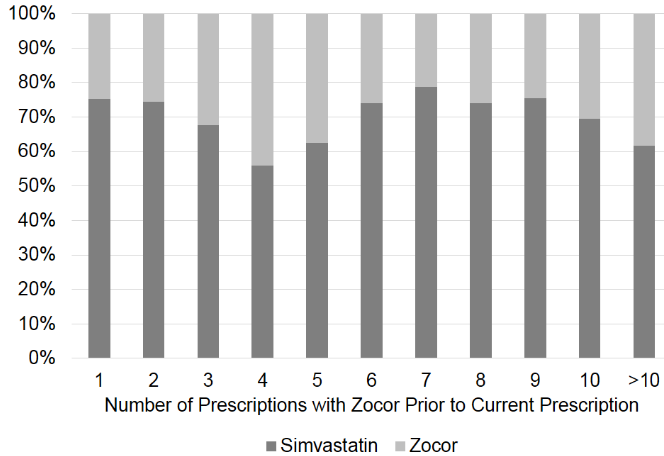


FIGURE 2.1: Choice Between Zocor and Simvastatin by Number of Prescriptions with Zocor

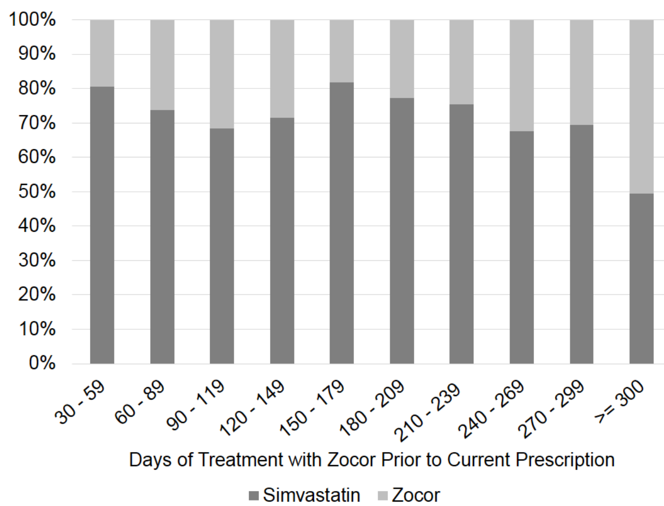


FIGURE 2.2: Choice Between Zocor and Simvastatin by Days of Treatment with Zocor

3

Competition and Dynamic Pricing in a Perishable Goods Market

3.1 Introduction

Sellers of perishable goods, such as airlines, ticket brokers, concert organizers and retailers of fashion and seasonal items, have to sell inventory within a fixed time horizon. These firms increasingly use dynamic pricing (DP) strategies, where they change prices as a function of both inventory and the time remaining, as technology makes it cheaper to change prices, track inventory and model consumer behavior. Managers often identify these types of revenue management strategies as being very valuable. For example, Robert Crandall, the former CEO of American Airlines, has been widely quoted as describing them as "the single most important technical development in transportation management since we entered the era of airline deregulation in 1979."[1]

[1] Smith et al. (1992) estimate that yield management increased AA's annual revenues by \$500 million. The San Francisco Giants implemented dynamic pricing for parts of their stadium in 2010 and estimated that it would increase their revenues by \$5m per year and the Giants' ticketing manager described DP as "changing the ticket world" (taken from an article by Adam Satarino in Bloomberg Businessweek, May 20 2010, accessed July 19, 2011).

The theoretical basis of DP strategies is now well-established, principally in cases where there is a single seller (see references below). However, there are few empirically-tractable models that can be used to guide managers' pricing decisions, especially in settings where (a) products are differentiated; (b) there is significant competition from other sellers; and (c) it is possible that the price that the manager sets will affect the prices that competitors set in the future. These features are present in all of the examples of perishable goods markets given above. In this paper, we introduce an empirical model that can be used for this purpose, and we show how it can be estimated using the type of data that a seller will typically have available. We use our model to analyze whether the pricing strategy that a seller currently uses is approximately optimal.

Our empirical setting is an online secondary ticket market (Stubhub) for sports event tickets. We use data for 15 home games played by a single major league sports team in 2010, where an anonymous large broker, who provided us with its sales data and whose explicit objective is revenue-maximization, accounted for a significant share of the market. In this market the broker faces competition from a large number of smaller sellers, many of whom are season ticket holders who do not want to attend a particular game. There are three features of the data that make the broker's dynamic price-setting problem an interesting one to study, both methodologically and substantively.

First, the broker cuts prices over time, as illustrated in Figure 3.1, where prices are relative to the face value of the ticket. The general pattern of falling prices has been documented using different data from secondary ticket markets by Sweeting (2012). The ability to reduce prices as the game approaches if tickets remain unsold makes it optimal to set higher prices further from the game, rationalizing a price-cutting pattern.

Second, most of the broker's sales are made in the last few days before the game when the broker's prices are lowest, as shown in Figure 3.2. This raises the question of whether it is optimal for the seller to delay so long before cutting prices, or whether he would be better off setting a slightly lower price further from the game in order to increase his probability of sale when competitors prices tend to be higher. It is not clear *ex-ante* whether the broker's current pricing rules achieve the trade-off between higher prices and lower probabilities of sale optimally.

Third, competition and price-setting work in an interesting way which is also found in lots of other markets but which has been ignored in the existing literature. Different sellers compete with each other, but they know that they update their prices only infrequently and, importantly, not at the same time. Therefore, when a seller sets his price today he knows that other sellers will treat this price as given when setting their prices in the future. We show below that non-brokers cut prices in response to the broker's price cutting policy, which may reduce the broker's profits, and our results suggest that the broker may respond sub-optimally to the prices set by non-brokers. From a modeling perspective, the existing literature has either modeled monopolists or a fixed number of competing firms who set prices simultaneously. Providing an empirical framework to model this type of stochastic pricing decisions is an important contribution of this paper, which may be relevant for thinking about a wide range of markets.

Central to our paper is an estimable continuous time model of the market where potential buyers, sellers and opportunities to change prices arrive stochastically. The optimal dynamic pricing policy for our broker depends qualitatively on several factors that are clearly related to particular parameters of the model: the arrival rate and preferences of consumers, and particularly their price sensitivity; the probability with which the broker can change prices and the number of listings that he has to sell; and, how the distribution of competitors' prices can be expected to respond to the price that the broker sets. We provide a quantitative assessment of the broker's optimal policy by estimating the parameters of the model and then performing counterfactual simulations. We find that, on average, the broker's prices close to the game are fairly close to optimal but that there are potential advantages to using different pricing rules further from the gain and from updating prices more frequently.

Our paper makes at least three contributions. First, we provide the first empirical framework for calculating optimal dynamic pricing policies in a setting with competing sellers, and we provide an assessment of a pricing policy that is used in this type of setting. Our framework could be applied to other markets where a seller has some ability to 'move the market' by affecting the distribution of prices that competitors set in the future. Second, we show how our model can be estimated using data that may be widely available to sellers. For example, we assume only the availability of transaction data from one seller, not competing sellers in the market. Our methods build on techniques recently developed for estimating continuous time games in the economics literature (Arcidiacono et al. (2012)). Finally, we can use the estimated model to highlight which parameters are key for determining the optimal pricing policy. We do this using counterfactual policies, where we vary the parameters from their estimated values.

We conclude this introduction by describing the most closely-related theoretical and empirical literature. Section 3.2 describes our data, and highlights the stylized facts that motivate our analysis. Section 3.3 presents the model and Section 3.4 describes how it is estimated. Section 3.5 presents the estimates and Section 3.6 the results of our counterfactuals. Section 3.7 concludes.

3.1.1 Literature Review

The theoretical literature on dynamic pricing, reviewed in Elmaghraby and Keskinocak (2003), dates back to Kincaid and Darling (1963). In this paper, and in later papers such as Gallego and Van Ryzin (1994), Bitran and Mondschein (1997) and McAfee and Velde (2008) a monopolist seller of a perishable good, with a fixed initial inventory, faces consumers who arrive stochastically, have to buy at once or exit the market and have unit demand with valuations drawn from a time-invariant distribution. In Gallego and Van Ryzin (1994) and McAfee and Velde (2008) the seller is assumed to update its price continuously, whereas Bitran and Mondschein (1997) assume that he updates it infrequently using a policy of periodic price reviews. With continuous updating, the optimal price is equal to the opportunity cost of a sale plus a mark-up, which depends on the shape of the valuation distribution (equivalently the flow demand curve). This opportunity cost will go down over time if sales are not made as future selling opportunities disappear. On the other hand, when a sale is made opportunity costs will increase. McAfee and Te Velde (2006) show that on average the seller tends to cut prices, if he has any units left, at the end of the time horizon.

Many of the assumptions made in these papers have been relaxed in the subsequent theoretical literature. Zhao and Zheng (2000) allow for time-varying elasticities of demand which can change how optimal prices tend to evolve. For example, if demand becomes more elastic closer to the game, which may happen if the type of consumer who visits the market changes, it will be optimal to set lower prices. We allow for this type of variation when we estimate our model. We also allow for competition between a small number of sellers. In the existing literature competition has been handled in a particular stylized way. For example, Gallego and Hu (2014) and Xu and Hopp (2006) provide methods for solving continuous time models with a fixed number of ex-ante symmetric firms. Lin and Sibdari (2009) solve a discrete time game where logit differentiated duopolists set prices each period, and they show how it can be optimal to raise its price in order to make it more likely that its competitor sells out and leaves the market. All of these models neglect the possibility that new sellers may enter, which is a clear feature of our data and allowed for in our model, and they also assume that sellers adjust prices simultaneously. This can mean either instantaneously in continuous time or every period in discrete time. Our model assumes that opportunities to change prices only arise stochastically, which is a realistic assumption for online markets where firms may only review their prices infrequently. If prices are strategic complements, the stochastic nature of price setting might increase incentives to raise prices further from the game. One feature which is not included in our model, which has been included in some of the recent theoretical literature, is strategic consumers, who can respond to higher prices by waiting in the market. This can make demand a long time before the game more elastic, leading to lower prices. Sweeting (2012) provides evidence against this type of strategic consumer behavior being of first-order importance in secondary market for event tickets.

The empirical dynamic pricing literature has focused on two questions. The first question is which existing theories of dynamic pricing explain how sellers price, assuming that they behave optimally (McAfee and Te Velde (2006), Puller et al. (2009), Sweeting (2012)). The second question is whether firms set prices optimally, assuming that a particular model is correct and this is also the question in our paper. The existing studies of this question, such as Heching et al. (2002), Vulcano et al. (2010), Caro and Gallien (2012), Soysal and Krishnamurthi (2012) and Li et al. (2011), focus on a single seller, although Vulcano et al. (2010) allow for consumers to substitute between multiple products sold by the same seller, in that case a fashion retailer. In our paper, we want to investigate how competition between sellers - in particular a large broker and many smaller sellers - should affect pricing. There are two elements to this competition effect: first, as other sellers cut prices, it may be optimal for the broker to set a lower price, and the importance of this may vary with the time until the game; and, second, the broker might strategically want to influence how other sellers prices evolve as suggested above. In doing so, we assume that consumers are not strategic (i.e., they have to buy as soon as they arrive or exit the market), an assumption that has been relaxed in Soysal and Krishnamurthi (2012) and Li et al. (2011), although including strategic consumers in the model would be an interesting direction for future research.[2]

Solving for equilibrium strategies in a dynamic pricing model with many, possibly asymmetric, competing sellers is beyond the current literature. In this paper we therefore take a simpler approach to modeling seller behavior, based on the estimation of policy functions as functions of the state variables. This approach has been widely used in the literature estimating dynamic games in economics (Bajari et al. (2007), Ryan (2012)), although applying it here requires us to innovate because our data involves repeated snapshots of a market that evolves continuously. We use the estimated functions describing how other sellers' prices evolve, to conduct counterfactuals, as in Benkard et al. (2010). This makes the strong assumption that a change in the broker's policy would not change the entry, exit and pricing *policies* of other sellers, although it may change the exact prices that they set. There are some ways to test this assumption and we will do so in future work.

3.2 Data

3.2.1 Sources

Our data comes from three sources: a professional ticket broker, Stubhub.com and the official website of the major league sports team whose games we use in our analysis. The broker provided us with its transaction data for games played by the team from December 2009 to April 2010. This data includes details of the exact seats (section, row, seat number) sold, the sale price, the time at which the

[2] As mentioned before, the results in Sweeting (2012) suggest that strategic consumer behavior is not too important in secondary ticket markets, possibly because the nominal gains to optimal timing decisions are quite small. Li et al. (2011) find that the proportion of strategic consumers in most markets for airline tickets, where the nominal gains are potentially larger because prices are higher, is not too large.

broker registered the sale in its internal system and the distribution channel through which the sale took place. It also lists the tickets that the broker had that were not eventually sold, although for this team the broker sells the vast majority of his listings. In this paper we use data from 14 games and we only use data on listing and sales on Stubhub, which was the outlet for most of the broker's listings as well as being the largest online secondary market for event tickets.

The Stubhub data was collected using a web scraper from the 'buy page' for each game. Our aim was to collect a snapshot of all tickets available on Stubhub about once every three hours, although in practice problems with the web scraper caused by Stubhub changing the presentation of its website, meant that the dataset is not complete, which is one reason why we pool games in estimation. This data contains the section, row, number of seats available and list price for each listing posted. The seller can also indicate whether a smaller number of tickets can be purchased, and sellers typically do this when they are trying to sell three or more seats. Stubhub does not show information about the seller as it provides a guarantee that someone buying from its site will receive tickets at least as good as those purchased. It also charges sizable commissions to both buyers (10%) and sellers (15%), and sets the shipping fees that have to be paid.

To estimate demand for listings on Stubhub, we need to match the listing data on Stubhub with the broker data. The aim of this matching process is to both identify which tickets on Stubhub belong to the broker and the prices that he set prior to sale, and to identify when broker sales occur and what was the competition that the broker's listings faced at the time of sale. The matching process is described in detail in the Appendix. Finally, the team's website was used to obtain the single game price (face value) of tickets in every section. During our sample, the team had a successful season, with realized attendances of at least 96% of stadium capacity for all of the 14 games that we use in estimation.

3.2.2 Summary Statistics

Table 3.1 shows how many downloads of data from Stubhub that we have for each game as the game approaches. The left-hand column shows the number of number of days prior to the game (so the top line is the day of the game), and each entry gives the number of downloads of data that we have. For example, when the number is 8, we have one download on average every three hours. For the first games in our sample, we lack data a long time before the game simply because these games are in January and we only started collecting data in December. For the later games, we have a fairly complete dataset for the month prior to the game, and when we interpret our results we will focus on this final month which is when the market is busiest and the most interesting price dynamics occur (Sweeting (2012)) shows that this is also true in the market for MLB tickets, using data for all teams).

Table 3.2 shows summary statistics for the broker's and competitors listings on Stubhub, both in the full sample and in the set of game-sections that we use in estimation. The main criteria for selecting these listings is that these are game-sections where the broker has at most one listing (e.g., one group of 6 seats in the same row). On average, competitors set prices that are close to face value whereas the broker sets lower prices. This pattern is consistent - in an optimal pricing framework - with the broker having no value to unsold tickets, whereas other sellers may be willing to attend the game themselves. This pattern is true for both the full sample and the estimation sample, although more expensive sections are over-represented in the estimation sample. The broker never has seats in the front row of a section, whereas about 7% of competitors' listings are for front-row seats. These listings may well be posted by season-ticket holders who are more likely to own front-row seats. The average number of seats in a broker's listing is 4.27, with a range from 1 to 23. However, 25% of these listings had two seats, 32.1% had four seats and 26% had five or six seats, with 90% having six seats or less. When we estimate the model we only use those observations where the broker has no more than six seats to sell.

Table 3.3 shows many seats the broker sells when a transaction takes place. The most common transaction involves a pair of seats, often from a larger listing. As we illustrate below, this can have an important impact on the optimal profile of prices as the optimal price increases when a sale of a subset of seats from a listing takes place. At the moment when we estimate the dynamic model we assume that all buyers who arrive in the market only want two seats, although when we estimate the demand parameters we are more flexible. Going forward we will try to incorporate the arrival of buyers wanting more than two seats into the model. A comparison of the number of observations for the broker and competitors in the estimation sample indicates that the broker has about 20% of listings in this sample.

3.2.3 Stylized Facts and Current Broker Pricing Behavior

We now describe how prices evolve over time, and what are the key facts about how the broker currently sells tickets that motivate our interest in his optimal pricing policy.

As prices and games are heterogeneous, we examine the dynamics of prices using a regression model

$$p_{it} = X_{it}\beta + D_t\alpha + FE_i + \varepsilon_{it}$$

where X_{it} are observed characteristics (e.g., the number of seats, the number of the row and measures of the performance of the home and away teams that may affect demand), D_t are dummies measuring the number of days prior to the game and FE_i are game-face value fixed effects. Prices (p_{it}) are the Stubhub list price divided by the face value of the ticket, although using the log of the nominal price gives broadly similar implications.

The coefficients are shown in Table 3.4, using different subsets of the data. In columns (1)-(4) observations are listings. The specifications differ in whether the sample comprises the broker's listings, competitors' listings, and whether or not the observations are from game-sections included in the estimation sample. Here we see that based on all of the data, non-brokers tend to maintain fairly constant prices just above face value (add together the price 0-2 days before the game and the relevant time coefficient) until the last few days before the game. Figure 3.3 also shows this average price path. On the other hand, the broker cuts prices more smoothly as a game approaches, dropping prices by 20% of face value in the last five days before the game and an additional 30% of face value in the 10 days before that. As can be seen in the Figure, the broker's average list prices are always below those of the average non-broker.

However, despite always having lower prices, most of the broker's sales are concentrated immediately before the game. Figure 3.4 shows how the sales rate of the broker evolves as a game approaches, where this rate is defined as the number of sales that the broker makes during the day divided by the number of listings that the broker has at the start of the day. This rate can be greater than 1 because a single listing can result in more than one transaction if multiple subsets of the tickets listed are sold. The sales rate climbs steeply in the last five days before the game. This could be caused by the broker's falling price, a decline in the number of competitor listings, which can be seen in Figure 3.5, or an increase in the number of buyers in the market. If the last two factors drive the increase in demand, this might suggest that the broker could raise his revenues by not cutting prices so dramatically. On the other hand, if it is the broker's lower price that increases his demand, this may suggest that the broker would do better by setting a slightly lower price earlier before the game to increase the likelihood that he sells. Our model helps us to distinguish what drives the increase in sale, and therefore to establish what the optimal pricing policy should be.

A comparison of columns (1) and (3) in Table 3.4 also highlights another potentially important factor in a seller's pricing decision. Using the full sample, which mainly consists of sections where the broker is not trying to sell tickets, competitors' maintain fairly constant prices until quite close to the game. On the other hand, when we use the estimation sample, which only includes sections where the broker has tickets and is cutting prices, we observe that non-brokers are cutting prices too. Assuming that it is the price cutting behavior of the broker that causes the non-brokers to cut their prices (which should be true as long as the sections in which the broker has tickets are random, as all of these specifications include game-face value fixed effects), this suggests that if market demand is low in the weeks prior to the game, the broker may be hurting himself when he cuts prices by causing his competitors' prices to fall, thereby reducing the price that he can set when the market is very active close to the game.

3.3 Model

Calculating the broker's optimal pricing rule requires a model that predicts (i) the probability that the broker will sell as a function of the broker's price and the price of competitors; and (ii) how competitors' prices will evolve as a function of the broker's price. We now describe the parsimonious, continuous time model that provides us with these predictions.

In outline, the model works in the following way: at any point in time the broker and his competitors have given numbers of tickets and particular prices for their listings. A number of stochastic events, with Poisson arrivals, can happen. These events are: (i) a potential buyer may arrive in the market, in which case she can choose to buy one of the available listings or exit the market forever. We therefore assume away the possibility that potential buyers can strategically delay purchasing, so that a buyer has a static, rather than a dynamic, problem to solve; (ii) the broker may change his price; (iii) a competitor may enter the market, exit the market (without being sold, e.g., the seller decides to go to the game himself) or change his price. We model a specific arrival rate for each of these events and we model, as a stochastic process, what happens when one of these events occurs.

To explain the model we begin by specifying the static, logit preferences of a buyer. This is useful as the state space of the dynamic model is defined using these preferences, following ideas in Nevo and Rossi (2008). We then explain how we parametrize the arrival rates and the evolution of the state variables. We model markets at the game-section level, and, in estimation, we use game-sections where the broker has only one set of tickets to sell (e.g., a group of 6 seats in row 8). In some ways this is too narrow, as consumers are likely to substitute across similar sections, but this approach means that we only have to deal with the broker setting a single price, which is easier in the counterfactuals. In future iterations, we hope to be more general. When estimating the dynamic part of the model we also assume that buyers are only interested in buying a pair of seats (recall, 54% of purchases are of two seats), although we are more flexible when estimating preferences.

3.3.1 Buyer Demand

We follow most of the literature on consumer demand by assuming that a buyer i arriving in the market will choose the listing that maximizes his utility, where his utility for listing j is given by a linear function:

$$u_{ij} = x_j\beta - \alpha p_j + \varepsilon_{ij} \equiv \delta_j + \varepsilon_{ij} \tag{3.1}$$

where x_j are listing j's characteristics (specifically, the row), p_j is the price (measured relative to face value) and ε_{ij} is a Type I extreme value (logit) error that independently and identically distributed across listings. The δs are known as "mean utilities". When faced by a set of K listings in a market, the probability that consumer buys listing j is

$$Pr(i \text{ chooses } j) = \frac{\exp(\delta_j)}{1 + \sum_{k=1}^{K} \exp(\delta_k)} \tag{3.2}$$

where the 1 in the denominator reflects the fact that the consumer may choose not to purchase any tickets. We assume that a consumer who wants to buy two seats chooses from the set of listings that have at least two seats (this implicitly assumes that all listings with more than two seats allow only two seats to be bought, which is true in 85% of cases, and in 95% of cases with four or more seats), whereas a consumer who wants to buy four seats chooses from the set of listings with at least four seats, and so on. Apart from through this effect on the choice set, the number of seats in a listing is assumed to have no effect on choice probabilities.[3]

3.3.2 State Space

Dealing with a large state space, containing information on the prices, rows and number of seats being sold in each listing is infeasible (at least using our current estimation methodology). Therefore we need to reduce the number of state variables, and to do so we base our model on the δ (mean utility) terms implied by preferences. Specifically, the state of a particular game-section is defined by the number of seats that the broker has left to sell (2, 4 or 6), δ_B of the broker's listing and the value of δ_{NB} for non-brokers' listings where

$$\delta_{NB} = \log\left(\sum_{k \in NB} \exp(\hat{\delta}_k)\right).$$

For estimation, we discretize the values of the δ_B and δ_{NB} state variables. Specifically we allow for 23 bins of δ_B, including an absorbing state associated with all of

[3] For example, this rules out the possibility that someone who wants to buy two seats might prefer to buy two out of four seats in order to increase the probability that they will have an empty seat next to them.

the broker's tickets having been sold, and 23 bins for δ_{NB}, including a state where no non-broker tickets are available. The boundaries of the bins are chosen so that roughly equal number of observations are in each bin.

3.3.3 State Transition Probabilities

In a continuous time, the probability of a transition from one state to another is determined by the arrival rate of the relevant event and the conditional probability of that particular transition given that an event arrives. For example, the probability that the number of broker's tickets falls by two is determined by the arrival rate of a buyer and the probability that an arriving buyer purchases from the broker's listing. Similarly, the probability that δ_B changes from one value to another is determined by the arrival rate of a move where the broker can change his price and the probability that the price change takes the value of δ_B to its new level. In continuous time we can treat these events as happening individually, as the probability that two events happen at the same instant is zero. I now describe each of these stochastic processes in turn.

The modeling of state transitions statistical processes is quite similar to approaches that have recently been taken to the estimation of dynamic games where players are assumed to use Markov Perfect equilibrium strategies (Ryan (2012); Benkard et al. (2010); Bajari et al. (2007)). However, although we assuming that transitions are functions of the previous state, we do not assume that either the broker or his competitors set prices in an optimal way. This reflects the fact that our question pre-supposes that even the largest seller may not set prices optimally.

Buyer Arrival and Purchase

Two broker tickets are sold if a buyer arrives and chooses to buy from the broker's listing. The Poisson arrival rate for a buyer is given by λ^{BUYER} and the probability that an arriving buyer purchases two tickets from the broker is

$$\frac{\exp(\delta_B)}{1+\exp(\delta_B)+(I_{NB}\neq 0)\exp(\delta_{NB})}$$

where $(I_{NB} \neq 0)$ means that there is a non-broker listing present. The arrival rate for buyers is allowed to vary as a game approaches, so that we can capture the market becoming more active over time. If a broker ticket is bought, the number of remaining broker tickets changes mechanically or, if the final broker tickets are bought, the state changes to the absorbing one where there are no broker tickets left.

Of course, it may be the case that the buyer arrives and buys non-broker tickets which may cause the value of δ_{NB} to change. The probability that a non-broker listing is sold is

$$\frac{\exp(\delta_{NB})}{1+\exp(\delta_B)+(I_{NB}\neq 0)\exp(\delta_{NB})}$$

and we model the evolution of δ_{NB} if this happens using a transition matrix P^{NB} which allows for the possibility for δ_{NB} to evolve to any lower value or stay the same. If the consumer chooses to buy no listing, he exits the market without the state changing.

Broker Price Change

We assume that opportunities for a price change arrive stochastically. The Poisson arrival rate for a broker change is λ^{BROKER} and, conditional on a move arriving, δ_B is assumed to evolve according to a stochastic process:

$$\delta_B^{a'} = \alpha_0 + \alpha_1\delta_B^a + \alpha_2(I_{NB}\neq 0)\cdot\delta_{NB}^b + \alpha_3(1-(I_{NB}\neq 0)) + \xi_1 \tag{3.3}$$

where ξ_B is assumed to be normally distributed with mean 0 and variance σ_1^2. Given the discrete nature of the state space this implies that the probability of the next value of δ_B being $\delta_B^{a'}$ is

$$\frac{\phi\left(\frac{\delta_B^{a'}-\alpha_0-\alpha_1\delta_B^a-\alpha_2(I_{NB}\neq 0)\cdot\delta_{NB}^b-\alpha_3(1-(I_{NB}\neq 0))}{\sigma_1}\right)}{\sum_{a''}\phi\left(\frac{\delta_B^{a''}-\alpha_0-\alpha_1\delta_B^a-\alpha_2(I_{NB}\neq 0)\cdot\delta_{NB}^b-\alpha_3(1-(I_{NB}\neq 0))}{\sigma_1}\right)} \tag{3.4}$$

where $\phi(\cdot)$ is the standard normal probability density function and the αs and σ_1 are parameters to estimate.

Non-Broker State Change

δ_{NB} may change if more sellers enter the market, sellers leave the market or an existing seller changes his price. As we have aggregated non-broker listings into a single variable, we cannot model individual decisions, but we do model three process by which δ_{NB} may change.

The first process, which one can think of as "non-broker arrival" allows for the possibility that non-broker listings will enter the market when there were none previously. The Poisson arrival rate is $\lambda^{NB,IN}$, and the probability that the state $(n_B, \delta_B^a, I_{NB} = 0)$ moves to state $(n_B, \delta_B^a, \delta_{NB}^b)$ is

$$\frac{\exp(\eta_1\delta_B^a + \eta_2\delta_{NB}^b + \eta_3\delta_B^a * \delta_{NB}^b)}{1 + \sum_{\delta_{NB}^k, k\neq 0}\exp(\eta_1\delta_B^a + \eta_2\delta_{NB}^k + \eta_3\delta_B^a * \delta_{NB}^k)} \tag{3.5}$$

where the ηs are parameters to estimate. Note that this function allows the new value of δ_{NB} to depend on the value of δ_B as we would expect it to do if competitors set lower prices when the broker has a lower price. We could enrich this specification to also depend on the number of tickets that the broker has left.

The second process, which one can think of as "non-broker exit" allows for the possibility that non-broker listings will exit the market without being sold. The Poisson arrival rate is $\lambda^{NB,OUT}$ and the probability that the state changes from

$(n_B, \delta_B^a, \delta_{NB}^b)$ to $(n_B, \delta_B^a, I_{NB} = 0)$ is

$$\frac{\exp(\kappa_1 \delta_B^a + \kappa_2 \delta_{NB}^b + \kappa_3 \delta_B^a * \delta_{NB}^b)}{1 + \exp(\kappa_1 \delta_B^a + \kappa_2 \delta_{NB}^b + \kappa_3 \delta_B^a * \delta_{NB}^b)} \tag{3.6}$$

where the κs are parameters to be estimated.

The third process is a change in δ_{NB} which does not involve the $I_{NB} = 0$ state. One can think of this a non-broker price change although it could also represent the entry or exit of a subset of non-broker listings. The Poisson arrival rate of this event is $\lambda^{NB,CHANGE}$, and we assume that δ_{NB}would evolve according to an stochastic process where

$$\delta_{NB}^{b'} = \tau_0 + \tau_1 \delta_B^a + \tau_2 \delta_{NB}^b + \xi_2 \tag{3.7}$$

where ξ_2 is assumed to be normally distributed with mean 0 and variance σ_2^2. The corresponding conditional probability can be expressed in a form similar to (3.4).

3.4 Estimation

There are two steps in our estimation. In the first stage we estimate the preference parameters, in order to define the value of the state variables. In the second stage we estimate the transition functions.

3.4.1 Preference Parameters

As discussed above, the probability that a consumer who arrives in the market purchases a ticket listed by the broker is

$$Pr(i \text{ chooses } j) = \frac{\exp(\delta_j)}{1 + \sum_{k=1}^{K} \exp(\delta_k)} \tag{3.8}$$

where

$$\delta_j = x_j \beta - \alpha p_j \tag{3.9}$$

To estimate the parameters β and αwe have to take into account that no buyers may arrive. For clarity, suppose that we are interested in whether two seats are sold from a listing j that has four seats between two particular downloads. We parametrize this probability in the following way

$$Pr(\text{two tickets purchased from } j) = \frac{\exp(W_g\gamma)}{1+\exp(W_g\gamma)} \frac{\exp(x_j\beta - \alpha p_j)}{1+\sum_{k=1}^{K}\exp(x_k\beta - \alpha p_k)} \quad (3.10)$$

where the first term is used to represent in a reduced-form way the probability that a buyer who wants two seats arrives. W contains dummies for the number of tickets being purchased (e.g., constant, 4 seats or 6 seats), dummies for 0 to 5, 6 to 10, 11 to 20 and more than 21 days before the game, to capture differences in the arrival rates, and weekend, working day time dummy and nighttime dummies to capture the possibility that arrival rates may also differ during the day. For a two seat purchase the sum over listings in the denominator of the second term would be over all listings with at least two seats. For a four seat listing it would be over listings with at least four seats. Estimation is done using Maximum Likelihood.

We note that this specification is not ideal. In particular, it assumes that at most one customer turns up between downloads and that the set of listings available to that customer is the set of listings available at the first download. These assumptions are not too unreasonable given that we are using downloads that are three hours apart but it could be improved upon.

3.4.2 State Space Definition

As described above, we use the estimated preference parameters to calculate the values of δ_B and δ_{NB} that define the state space. The bins that we use for these values are shown in Figure 3.6.

3.4.3 Estimation of the Continuous Time Processes

Our data comes in the form of snapshots of the state of the Stubhub market every few hours, plus indicators for when a sale from a broker listing is made between the downloads. Our model of how the market evolves is in continuous time, so that any number of events may happen between downloads. We therefore need to use our model to calculate the probability that the market will be in each possible state at the next download given the current state. The math of continuous time Markov processes makes this straightforward to do.

The first step is to construct the "intensity matrices" (Q, dimension $1,519$ x $1,519$ where 1,519 is the total number of states) for each event, which summarizes the finite Markov jump process. Entry $Q_x(m,n)$ ($m \neq n$) equals the Poisson arrival rate of the event multiplied by the conditional probability of moving from state m to state n given that the relevant event arrives. The diagonal elements of the intensity matrix is the minus sum of the other elements in the row so that the sum of the elements in a row is 0. For example, the intensity matrix for the event associated with a buyer arriving, Q^{BUYER}, has the elements corresponding to the broker sale in state $(n_B, \delta^a_B, \delta^b_{NB})$ equal to arrival rate λ^{BUYER} multiplied by the probabilities associated with the broker's tickets being bought so that n_B falls by two, or non-broker tickets being purchased so that δ_{NB} changes according to matrix P^{NB}. Off-diagonal elements associated with the values of n_B or δ_{NB} increasing would be equal to zero. Intensity matrices for broker price changes (Q^{BROKER}), non-broker entry ($Q^{NB,IN}$), non-broker exit ($Q^{NB,OUT}$) and non-broker price change ($Q^{NB,CHANGE}$) are defined similarly.

The "aggregate intensity matrix" is calculated by summing up the intensity matrices, i.e. $Q = Q^{BUYER} + Q^{BROKER} + Q^{NB,IN} + Q^{NB,OUT} + Q^{NB,CHANGE}$. The transition matrix $P(t)$, which reflects the probability of transitioning from one state to the other after a time of length t (via any combination of state changes), can be found as the unique solution to the system of ordinary differential equations

$$P'(t) = P(t)Q$$
$$P(0) = I \tag{3.11}$$

where I is the identity matrix of the same size as $P(t)$. Solving the above system of equations gives $P(t) = e^{tQ}$, a matrix exponential which is calculated using EXPOKIT in MATLAB. With the transition matrix, we are able to find the corresponding likelihood for each observation and estimate the transition parameters using MLE. The log-likelihood function for transition parameters (θ_2) is

$$L(\theta_2) = \sum_{g} \sum_{j \in g} \log P(t_j, \Theta)(s_{j-1}, s_j) \tag{3.12}$$

where s_j is the state index for observation j, Θ are all of the parameters, g denotes the game, $j-1$ denotes the previous download and t_j is the time between the downloads of observation $j-1$ and j

3.5 Parameter Estimates

We now discuss the parameter estimates, with the counterfactuals presented in the next section.

Table 3.5 shows the first stage estimates. Higher prices significantly reduce the probability that a listing is sold, with an average own price elasticity, conditional on the arrival of a buyer, of -4.44, which is the range usually considered for consumer markets although it is below the elasticities sometimes found for undifferentiated products sold via price search engines on the internet (Ellison and Ellison (2009)).[4]

[4] Ellison and Ellison look at unbranded memory chips and other computer components sold by small firms that are listed on Pricewatch.com. For the lowest quality chips estimated elasticities are around -25, which is exceptionally high, whereas for medium and high quality chips, for which it is harder to compare prices, elasticities are in the range of 4 to 7. Obviously if there is some endogeneity of price (caused,by, for example, positive demand shocks) we may underestimate the elasticity of demand resulting in optimal prices that are too high. Future versions will include more fixed effects to see if we can get this elasticity to increase.

We also tried interacting the price coefficient with the number of days to go, to see if there was evidence consistent with the type of consumer who is in the market changing over time. However, in most of these specifications there was no clear pattern to the interaction coefficients and they were imprecisely estimated. The coefficient on row is negative, and significant at the 5% level, implying (sensibly) that consumers value rows that are further back less. We have tried to incorporate additional terms to capture whether front row seats are much more attractive, but the coefficient on this term was never precisely estimated. This partly reflects the fact that the broker never has front row seats in our sample. The coefficients on the W variables are also sensible, consistent with the market being more active close to the game, two seat buyers being the most common and the market being more active during the workday and at night than during the early morning (12 am to 6 am is the excluded category).

The second stage estimates of the state transitions are in Table 3.6. Before we discuss the results, it is worth commenting that one potential problem when estimating such a large number of parameters is that maximum likelihood may not find a global maximum. However, when we re-started the estimation from quite different sets of parameters we got almost identical coefficients, indicating that this is not a significant problem in our setting.

The signs and size of most of the coefficients are intuitive. Beginning with the arrival rates (λ), we see that all the arrival rates increase as a game approaches, consistent that there are both more consumers in the market and that sellers are more likely to change their prices as the moment when the tickets "perish" approaches. The high rate of non-broker exit in the last five days is also driven by Stubhub's rule that hard copies of tickets can only be sold in the last three days if they are provided to Stubhub. This constraint does not affect the broker, but is likely to impact small sellers.

The transition coefficients imply several intuitive effects. A higher value of δ_B, for example a **lower** broker price, implies that δ_{NB} is more likely to increase, i.e., that non-brokers are likely to lower their prices, and that if non-broker listings enter they will do so at higher δ_{NB}. These effects should create an incentive for the broker to keep his price high when the market is inactive. On the other hand, there are effects that offset this incentive. For example, a lower broker price makes it more likely that non-broker listings will exit, raising the broker's future demand, and creating an incentive for the broker to lower his price early on in order to drive other listings out. δ_{NB} also affects the broker's pricing rule in an intuitive way, with lower non-broker prices making it more likely that the broker will set a lower price.

3.6 Counterfactuals: Optimal Prices

This section describes the counterfactuals that we have run, calculating the broker's optimal price under various scenarios. Our focus at the moment is to compare the level and time profile of the broker's current prices with those which our model predicts should be optimal. We can also investigate how the broker's profits and optimal policies vary with the model's parameters. This can help to tell us more about the economics of the pricing problem, but it can also provide useful advice to the broker. For example, we can consider how valuable it is to change price more frequently.

It is important to note that when we consider the broker changing his price we allow for non-brokers to change their prices, or possibly enter or exit the market, according to these transitions. However, we assume that the value of the parameters of these functions would not change, i.e., we do not try to re-compute the full equilibrium of the pricing model. Our approach to performing counterfactuals is therefore similar to the one used in recent papers such as Benkard et al. (2010).

3.6.1 Calculations in a Base Case

We now explain how the broker's optimal price is calculated using the simplest example where the broker has two tickets to sell and his listing's ROW=10. We assume that the broker can change his price once each day, at the start of the day. For any value of this price, call it p_B, we calculate expected revenues in the following steps, starting one day before the game when the price set is p_B^{T-1} and the initial non-broker state is δ_{NB}^{T-1}:

(i) calculate the value of δ_B associated with $p_B^{T-1}$and ROW=10;

(ii) calculate the probability that the listing is sold within the next 24 hours using the appropriate transition matrix, $P(24)$ and the initial state;

(iii) calculate the value of the expected payoff to $p_B^{T-1}$where the payoff is equal to

$$p_B^{T-1}\,\mathrm{Pr}(sale \mid \delta_B(p_B^{T-1}), \delta_{NB}^{T-1}, \Theta) \tag{3.13}$$

which implicitly assumes that the broker has no value from unsold tickets and Θ are all of the parameters estimated above;

(iv) for each state δ_{NB} find the optimal price, i.e., the price that maximizes the payoff. The associated payoff is the "value" $V^{T-1}(\delta_{NB})$ to being in state $\delta_{NB}^{T-1}$with one day remaining;

(v) repeat the calculations in (i)-(iv) for the previous day $(T-2)$ where the expected payoff is now

$$\begin{aligned} &p_B^{T-2} Pr\Big(sale\,at\,T-2 \mid \delta(p_B^{T-2}), \delta_{NB}^{T-2}, \Theta\Big) + \\ &\quad \sum_k V^{T-1}(\delta^{T-1}(k)) Pr\Big(\delta_{NB}^{T-1}(k)\,and\,no\,sale\,in\,T-2 \mid \delta_B(p_B^{T-2}), \delta_{NB}^{T-2}, \Theta\Big) \end{aligned} \tag{3.14}$$

so the broker takes into account how the price that he sets at $T-2$ may influence the transition of δ_{NB}, which he cares about if he does not sell;

(vi) repeat (v) for $T-3$ and so until $T-30$.

This calculation assumes that the broker started with only two tickets to sell, so that the game ends as soon as a single sale is made. However, the logic extends quite naturally when we consider the broker having more tickets to sell. Suppose, for example, that the broker has 4 seats to sell two days before the game. In this case, he needs to take into account that, for a given price, he may sell all of his tickets today (to two different buyers given his assumptions), or he may sell two tickets, and so be left with two tickets to sell on the final day or he may sell no tickets and have four to sell on the final day. This simply involves adding additional terms to the calculation in 3.14.

3.6.2 Results

Base Case

Tables 3.7-3.9 show the optimal prices for each (day-to-game, non-broker state, number of tickets remaining) combination where we allow for the broker to start with up to six tickets. The top line of each table shows the average price for that day, averaging across non-broker states. Figure 3.7 compares these 'average state' paths to the evolution of actual prices (based on the regression results in Table 3.4 where we assume that the broker's listings are in row 10).[5]

At least four features of these results deserve discussion.

(i) the broker sets prices that are of approximately the correct level immediately before the game, which is when the market is most active. At this point, the pricing problem is fairly close to being a static one (as future opportunities to sell are limited), and finding that the broker's pricing decisions are consistent with our model when the pricing problem is static provides some indirect evidence that our demand

[5] We have also simulated what the average price would be given how the distribution of non-broker states can vary over time, and given the ability of the broker to influence how the non-broker state evolves. The differences in the average optimal prices are small, so we report average prices here.

estimates, which determine the optimal static pricing policy, are approximately correct. It is also worth noting that the average level of prices around one month before the game is also approximately correct for listings with 2 or 4 seats.

(ii) immediately before the game the optimal price are sensitive to the level of competition δ_{NB}, as can be seen in Tables 3.7-3.9, with stronger competition implying a lower optimal price, as one would expect in a static demand model. However, this sensitivity declines as one moves further from the game, so that more than 10 days before the game the optimal price is almost invariant to the non-broker state.[6] This reflects the fact that the broker knows that he is less likely to sell further from the game and that it is possible that the non-broker state will improve over time, and that by allowing low priced non-broker tickets to be sold, by setting a high price, the broker may be able to encourage this.

We can compare these predictions with how the broker currently prices. In Table 3.10 we repeat the regression reported in column (2) of Table 3.4, which used broker listings, adding two variables. The first one measures the minimum price of non-broker listings in the game-section and the second is a dummy variable that is equal to one when there are no non-broker listings. In second column of Table 3.10 these variables are interacted with a count of many days there are until the game to test whether the broker's prices are more or less sensitive to these measures of competition as a game approaches.

The positive coefficients on these variables in the first column indicate that the broker sets higher prices when he faces less competition. The effects are large relative to what an optimal pricing strategy would predict. For example, the 25th and 75th percentiles of the minimum price variable are 0.54 and 0.80, so that moving between these two values would predict that the broker's relative price should increase by 0.08.

[6] We note that the highest δ_{NB} is an obvious exception to this statement and we are in the process of figuring out why.

Optimal prices only display anything comparable to this sensitivity in the last couple of days before the game, and not on average. Similarly no non-broker competition implies a price increase of almost 0.25, which is a much larger effect than we see for optimal prices.

In the second column these additional variables are interacted with the number of days to go before the game. The coefficients on the main effects are zero, indicating insignificant differences immediately before the game, whereas optimal pricing implies large differences. In contrast, the positive coefficients on the interactions indicate that the prices are sensitive further from the game, and once again this is the opposite of what the optimal pricing results predict.

(iii) the optimal pricing policy involves setting lower prices when there are more tickets to sell as multiple purchases will be required to clear the inventory. This reflects a standard feature of dynamic pricing models (e.g., McAfee and Te Velde (2006)) where the opportunity cost of sale of a given seat decreases with the number of seats left to sell, which causes the optimal price to fall.In contrast, the broker sets similar current prices for listings with 2, 4 or 6 seats. This suggests that there may be gains (which we could quantify) to raising prices are subsets of seats from a listing are sold. One caveat here is that our assumption that buyers only ever want to purchase two seats probably leads to larger differences in optimal prices. In future revisions we will allow for some buyers to want four seats, and it will be interesting to see how much the differences in optimal prices fall;

(iv) while actual and optimal prices both fall as a game approaches the shape of the declines is not the same. Optimal prices have a concave time profile, which is exactly dynamic pricing models usually predict even in the absence of competition and intertemporal changes in market activity (see, for example, the simulations in McAfee and Te Velde (2006)). Adding these factors should make a concave (i.e., quickening pace of decline with initially fairly flat prices) even more optimal, as the optimizing seller will place more value on trying to keep competitors' prices high (by not cutting his own price) and less weight on trying to sell tickets early on by cutting prices because buyers are scarce. In contrast, the broker's actual prices decline fairly steadily starting about two weeks before the game, when we estimate that there are still relatively few consumers in the market, so that even with the price cut the probability of sale is low. We also fail to explain the small jump in the broker's prices that is observed about two weeks before the game. This is hard to explain in our model without introducing some additional element such as a change in the price elasticity of buyers' who enter the market during this period of time.[7]

Frequency of Price Setting

In future revisions we will vary many of the parameters to investigate how they affect optimal pricing and the relationship between optimal prices and the prices that we observe in the data. Here we investigate the role of how frequently the broker sets prices, holding the remaining parameters fixed.

The base case assumed that the broker updates prices once per day which is more frequently than we estimate happens in the data. To be precise our estimates imply that the probability that the broker updates his prices at least once during the day is 0.73 for 0-5 days before the game, 0.41 5-10 days before, 0.26 11-20 days before and 0.21 more than 21 days before the game. Figure 3.8 shows how optimal prices change when we assume that the broker updates with these probabilities, with the λ_B prices indicating the price that the broker should set if he gets the opportunity to change his price, knowing that he may not be able to change his price in the future. The key finding is that the when price changes are less frequent the broker should

[7] Interestingly when we estimated a more flexible demand model that allow for the price coefficient to vary with the number of days until the game we did estimate that demand was less elastic with two weeks ago than either more than 20 days or less than 10 days before the game, although the difference in the coefficients was not statistically significant.

set lower prices because there is some probability that the broker will be stuck with his earlier prices when the (unconstrained) optimal price falls. On the other hand, updating close to the game is sufficiently likely that the differences in optimal prices in the last few days are actually quite small.

3.7 Conclusion

In this paper we provide an empirical framework for analyzing settings where a seller faces both a dynamic pricing problem and significant competition, and is able to influence how his competition evolves when he sets his price. We estimate our model using data from a large broker who sells event tickets on Stubhub. In the data we see that the prices that the broker sets appear to significantly affect the prices of his competitors, as well as affecting the probability with which his tickets are sold. We find evidence that the broker's current pricing rules are optimal in some respects but not in others. For example, close to the game, when the market is most active, the broker sets prices that are approximately the right level to maximize revenues. On the other hand, the broker's prices further from the game are (i) too insensitive to the number of tickets that the broker has left; (ii) too sensitive to the current level of competition that the broker faces; and (iii) updated less frequently than what would have been optimal. In future revisions we flesh out these differences in more detail.

Table 3.1: Number of Successful Downloads Each Day For Each Game (more than 8 downloads listed as 8, days more than 45 prior to game not listed)

Days Prior	**Game**													
to Game	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**
0	7	6	6	0	1	6	8	8	8	8	8	5	8	8
1	8	8	8	0	1	8	8	8	8	8	8	8	8	8
2	8	8	8	4	1	6	8	8	8	8	8	8	8	8
3	8	8	8	8	0	1	8	8	8	8	8	8	8	8
4	8	8	8	8	3	1	8	8	8	8	8	8	8	8
5	1	8	8	8	8	1	8	8	8	8	8	8	8	8
6	3	8	8	8	8	1	6	8	8	8	8	8	8	8
7	8	1	8	8	8	1	1	8	8	8	8	8	8	8
8	8	3	8	8	8	1	1	6	8	8	8	6	8	8
9	8	8	1	8	8	0	1	1	8	8	8	7	8	8
10	8	8	3	8	8	3	1	1	6	8	8	8	8	8
11	8	8	8	8	8	8	1	1	1	8	8	8	8	8
12	8	8	8	8	8	8	1	1	1	8	8	8	8	8
13	6	8	8	8	8	8	0	1	1	8	8	8	8	8
14	5	8	8	1	8	8	3	1	1	8	8	8	8	8
15	4	5	8	3	8	8	8	0	1	6	8	7	8	8
16	8	6	8	8	8	8	8	3	1	1	8	8	8	8
17	8	4	5	8	8	8	8	8	0	1	8	8	8	8
18	5	8	6	8	8	8	8	8	3	1	8	8	8	8
19	5	8	4	8	6	8	8	8	8	1	8	8	8	8
20	5	3	8	8	7	8	8	8	8	1	6	7	8	8
21	1	2	8	8	8	8	8	8	8	1	1	8	8	8
22	3	3	3	5	8	8	8	8	8	0	1	4	8	8
23	1	1	2	6	8	8	8	8	8	3	1	1	8	8
24	2	2	1	4	7	8	8	8	8	8	1	1	8	8
25	0	1	1	8	8	7	8	8	8	8	1	1	8	8
26	0	2	2	8	8	7	8	8	8	8	1	1	8	8
27	0	0	1	8	8	8	8	8	8	8	0	1	6	8
28	0	0	2	8	8	8	8	8	8	8	3	1	1	8
29	0	0	0	8	8	8	6	8	8	8	8	0	1	6
30	0	0	0	0	5	7	7	8	8	8	7	0	1	1
31	0	0	0	2	0	8	8	7	8	8	8	0	1	1
32	0	0	0	1	0	8	8	7	8	8	8	0	1	1
33	0	0	0	2	0	8	8	8	6	8	8	0	1	1
34	0	0	0	0	0	8	7	8	8	8	8	5	0	1
35	0	0	0	0	0	8	8	8	8	8	8	8	3	1
36	0	0	0	0	0	5	8	7	8	8	8	8	8	0
37	0	0	0	0	0	0	8	8	8	8	8	8	8	3
38	0	0	0	0	0	0	8	8	7	6	8	8	8	8
39	0	0	0	0	0	0	8	8	8	8	8	8	8	8
40	0	0	0	0	0	0	5	8	8	8	8	8	8	8
41	0	0	0	0	0	0	0	8	8	8	8	8	8	8
42	0	0	0	0	0	0	0	5	8	8	8	8	8	8
43	0	0	0	0	0	0	0	0	8	7	7	8	8	8
44	0	0	0	0	0	0	0	0	5	8	7	8	8	8
45	0	0	0	0	0	0	0	0	0	8	8	6	8	8
No. of Broker Sales	66	50	66	65	116	149	111	69	150	59	76	37	81	77

Table 3.2: Summary Statistics

	Full Sample					Sample for Estimation				
Variable	Obs	Mean	Std. Dev.	Min	Max	Obs	Mean	Std. Dev.	Min	Max
Non-Broker Listings										
Face Value ($ per seat)	1,455,198	78.21	39.25	29.50	127.00	127,795	97.86	18.07	62.00	127.00
List Price ($ per seat)	1,455,198	78.82	49.07	1.95	1400.00	127,795	97.70	42.58	15.00	433.00
List Price Relative to Face (%)	1,455,198	1.14	0.91	0.07	47.46	127,795	1.03	0.51	0.12	4.46
Number of Seats Available	1,455,198	3.48	2.26	1.00	30.00	127,795	3.27	1.33	2.00	6.00
Front Row in Sections	1,455,198	0.07	0.25	0.00	1.00	127,795	0.05	0.21	0.00	1.00
Days Prior to Game Listing Present	1,455,198	34.51	26.08	0.00	101.00	127,795	39.55	27.76	0.00	101.00
Broker Listings										
Face Value ($ per seat)	87,401	94.63	21.83	62.00	127.00	32,361	92.45	21.68	62.00	127.00
List Price ($ per seat)	87,401	71.46	33.43	7.00	245.00	32,361	75.24	37.28	7.00	245.00
List Price Relative to Face (%)	87,401	0.78	0.34	0.11	1.93	32,361	0.82	0.36	0.11	1.93
Number of Seats Available	87,401	4.27	2.11	1.00	23.00	32,361	3.25	1.21	2.00	6.00
Front Row in Sections	87,401	0.00	0.00	0.00	0.00	32,361	0.00	0.00	0.00	0.00
Days Prior to Game Listing Present	87,401	37.47	28.64	0.00	101.00	32,361	39.82	27.62	0.00	101.00

Table 3.3: Number of Seats in Broker Sale

Number of Seats	**Transactions**	**% of Transactions**
1	45	2.57
2	940	53.59
3	284	16.19
4	332	18.93
5	68	3.88
6	56	3.19
7	10	0.57
8	14	0.80
9	2	0.11
10	3	0.17
Total	1,754	100

Table 3.4: Price Regressions

	(1) Non-Broker Listings	(2) Broker Listings	(3) Non-Broker Listings	(4) Broker Listings	(5) Broker Transactions
Sample	Full	Full	Estimation	Estimation	Full
Ave Price 0 to 2 Days Before Game	0.84	0.47	0.64	0.46	0.41
Days-to-go coefficients					
2 to 4 days	0.305***	0.0975***	0.168***	0.106***	0.0590***
	(0.033)	(0.023)	(0.018)	(0.020)	0.014
5 to 7 days	0.292***	0.216***	0.239***	0.235***	0.107***
	(0.031)	(0.028)	(0.020)	(0.027)	0.016
8 to 10 days	0.284***	0.342***	0.276***	0.362***	0.118***
	(0.032)	(0.027)	(0.023)	(0.026)	0.02
11 to 13 days	0.304***	0.498***	0.294***	0.517***	0.220***
	(0.033)	(0.024)	(0.023)	(0.025)	0.029
14 to 16 days	0.309***	0.519***	0.305***	0.528***	0.203***
	(0.030)	(0.024)	(0.022)	(0.024)	0.036
17 to 19 days	0.300***	0.401***	0.304***	0.387***	0.217***
	(0.030)	(0.026)	(0.022)	(0.026)	0.025
20 to 22 days	0.320***	0.442***	0.323***	0.408***	0.223***
	(0.033)	(0.029)	(0.023)	(0.027)	0.027
23 to 28 days	0.291***	0.458***	0.333***	0.427***	0.177***
	(0.034)	(0.027)	(0.023)	(0.025)	0.024
29 to 34 days	0.291***	0.491***	0.337***	0.451***	0.154***
	(0.034)	(0.025)	(0.022)	(0.024)	0.029
35 or more days	0.305***	0.560***	0.411***	0.534***	0.112***
	(0.029)	(0.023)	(0.021)	(0.023)	0.028
Number of seats (2 = excluded category)					
1	-0.315***	-0.407***	-0.251***	-0.594***	-0.305***
	(0.035)	(0.057)	(0.043)	(0.028)	0.11
3	0.304***	0.0412*	-0.0233	0.0288	-0.00889
	(0.058)	(0.024)	(0.025)	(0.030)	0.015
4	0.213***	-0.0028	0.0464*	0.000462	0.00391
	(0.019)	(0.015)	(0.025)	(0.018)	0.0086
5	0.405***	0.0153	-0.00565	0.0000518	0.0162
	(0.050)	(0.017)	(0.027)	(0.022)	0.012
6	0.262***	0.0409*	-0.125***	0.0514**	0.00397
	(0.023)	(0.021)	(0.030)	(0.026)	0.0099
7+	0.325***	0.0111	-0.106***	-0.0402	0.0102
	(0.057)	(0.020)	(0.028)	(0.028)	0.0093
Team performance (league rank)					
Home	0.0114***	0.0553***	0.00584**	0.0563***	0.0123***
	(0.004)	(0.002)	(0.002)	(0.003)	0.0028
Away	-0.00172	0.00881***	0.000798	0.0038	-0.00694**
	(0.006)	(0.003)	(0.003)	(0.004)	0.0033
Game-face value fixed effects	Y	Y	Y	Y	Y
Row Controls	Y	Y	Y	Y	Y
Observations	1,455,198	87,401	244,896	53,022	1172
R-squared	0.21	0.7	0.59	0.72	0.71

Note: standard errors in parentheses robust to heteroskedasticity and clustered on the game. ***,** and * denote statistical significance at the 1, 5 and 10% levels.

Table 3.5: First Stage Estimates

	Parameter	Std. Error
Demand (X)		
Constant	3.4621	(0.7870)
Price (relative to face)	-6.1270	(0.6833)
Row	-0.0507	(0.0205)
Arrival (W)		
Constant	-5.6897	(0.2855)
I(4 ticket people)	-0.3998	(0.1206)
I(6 ticket people)	-0.8741	(0.2783)
DTG0to5	2.9138	(0.1613)
DTG6to10	1.8241	(0.2069)
DTG11to20	1.4052	(0.2000)
DTG21to30	0.8094	(0.1968)
Weekend	-0.2122	(0.1106)
Daytime	1.7024	(0.1952)
Nighttime	1.7297	(0.2087)

Table 3.6: Second Stage Estimates

	Parameter	Std. Error		Parameter	Std. Error
λ - B move			NB transition		
0-5 days to go	0.0551	(0.0035)	Constant	0.4923	(0.0501)
6-10	0.0220	(0.0021)	B state	0.1024	(0.0243)
11-20	0.0127	(0.0010)	NB state	1.0143	(0.0334)
21-30	0.0097	(0.0009)	sigma	0.6113	(0.0396)
λ - NB move			B transition		
0-5 days to go	0.0676	(0.0061)	Constant	1.1978	(1.0522)
6-10	0.0272	(0.0031)	B state	1.0612	(0.3495)
11-20	0.0208	(0.0018)	NB state	1.0271	(0.2985)
21-30	0.0171	(0.0016)	I(No NB listing)	3.6013	(1.2180)
			sigma	3.9054	(0.5728)
λ - NB exit					
0-5 days to go	0.0331	(0.0062)	NB exiting		
6-10	0.0104	(0.0042)	B state	1.7093	(0.4335)
11-20	0.0058	(0.0021)	NB state	-1.0185	(0.3702)
21-30	0.0029	(0.0009)	Bstate*NBstate	-0.1089	(0.0762)
λ - NB enter			NB entering		
0-5 days to go	0.0539	(0.0069)	B state	1.5853	(0.2235)
6-10	0.0131	(0.0048)	NB state	-0.0348	(0.0355)
11-20	0.0068	(0.0023)	Bstate*NBstate	0.0534	(0.0207)
21-30	0.0032	(0.0016)			
λ - Consumer					
0-5 days to go	0.2326	(0.0114)			
6-10	0.0714	(0.0073)			
11-20	0.0431	(0.0039)			
21-30	0.0363	(0.0030)			

Table 3.7: Optimal Pricing with Daily Price Changes (Six Seats)

		Days Prior to Game																													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
	Mean price	**0.46**	**0.58**	**0.67**	**0.72**	**0.77**	**0.82**	**0.83**	**0.84**	**0.85**	**0.86**	**0.86**	**0.87**	**0.87**	**0.88**	**0.88**	**0.88**	**0.89**	**0.89**	**0.90**	**0.90**	**0.91**	**0.91**	**0.91**	**0.92**	**0.92**	**0.92**	**0.92**	**0.93**	**0.93**	**0.93**
	Non NB tickets	0.49	0.62	0.70	0.76	0.81	0.81	0.82	0.84	0.85	0.86	0.86	0.86	0.87	0.88	0.88	0.89	0.89	0.90	0.90	0.91	0.89	0.90	0.90	0.91	0.91	0.91	0.92	0.92	0.92	0.93
	-12.83	0.50	0.63	0.70	0.76	0.80	0.80	0.81	0.82	0.83	0.84	0.84	0.84	0.85	0.85	0.86	0.86	0.87	0.87	0.88	0.88	0.88	0.89	0.89	0.89	0.90	0.90	0.90	0.91	0.91	0.91
	-4.60	0.50	0.63	0.71	0.77	0.81	0.82	0.83	0.84	0.85	0.85	0.86	0.86	0.87	0.87	0.87	0.88	0.88	0.88	0.89	0.89	0.89	0.89	0.90	0.90	0.90	0.90	0.91	0.91	0.91	0.91
	-3.32	0.49	0.62	0.70	0.75	0.79	0.80	0.81	0.81	0.82	0.83	0.83	0.84	0.84	0.84	0.85	0.85	0.85	0.86	0.86	0.86	0.87	0.88	0.88	0.88	0.88	0.89	0.89	0.89	0.89	0.90
	-2.50	0.48	0.61	0.69	0.74	0.78	0.79	0.79	0.80	0.81	0.82	0.82	0.83	0.83	0.83	0.84	0.84	0.85	0.85	0.85	0.86	0.87	0.87	0.87	0.87	0.88	0.88	0.88	0.88	0.89	0.89
	-2.16	0.48	0.61	0.69	0.74	0.78	0.78	0.79	0.80	0.81	0.82	0.82	0.82	0.83	0.83	0.84	0.84	0.84	0.85	0.85	0.86	0.86	0.87	0.87	0.87	0.88	0.88	0.88	0.88	0.89	0.89
	-1.94	0.48	0.61	0.68	0.74	0.78	0.78	0.79	0.80	0.81	0.81	0.82	0.82	0.83	0.83	0.84	0.84	0.84	0.85	0.85	0.86	0.86	0.87	0.87	0.87	0.87	0.88	0.88	0.88	0.89	0.89
	-1.71	0.48	0.60	0.68	0.73	0.77	0.78	0.79	0.79	0.80	0.81	0.82	0.82	0.83	0.83	0.83	0.84	0.84	0.85	0.85	0.85	0.86	0.86	0.87	0.87	0.87	0.88	0.88	0.88	0.89	0.89
	-1.51	0.47	0.60	0.68	0.73	0.77	0.77	0.78	0.79	0.80	0.81	0.82	0.82	0.82	0.83	0.83	0.84	0.84	0.85	0.85	0.85	0.86	0.86	0.87	0.87	0.87	0.88	0.88	0.88	0.89	0.89
δ_{NB}	-1.33	0.47	0.60	0.67	0.73	0.77	0.77	0.78	0.79	0.80	0.81	0.81	0.82	0.82	0.83	0.83	0.84	0.84	0.84	0.85	0.85	0.86	0.86	0.87	0.87	0.87	0.88	0.88	0.88	0.88	0.89
	-1.15	0.47	0.59	0.67	0.72	0.76	0.77	0.78	0.79	0.80	0.81	0.81	0.82	0.82	0.83	0.83	0.83	0.84	0.84	0.85	0.85	0.86	0.86	0.87	0.87	0.87	0.87	0.88	0.88	0.88	0.89
	-0.99	0.46	0.59	0.67	0.72	0.76	0.77	0.78	0.79	0.80	0.81	0.81	0.82	0.82	0.83	0.83	0.83	0.84	0.84	0.85	0.85	0.86	0.86	0.86	0.87	0.87	0.87	0.88	0.88	0.88	0.89
	-0.86	0.46	0.58	0.66	0.72	0.76	0.77	0.78	0.79	0.79	0.80	0.81	0.82	0.82	0.82	0.83	0.83	0.84	0.84	0.85	0.85	0.86	0.86	0.86	0.87	0.87	0.87	0.88	0.88	0.88	0.89
	-0.71	0.46	0.58	0.66	0.72	0.76	0.77	0.78	0.79	0.80	0.80	0.81	0.82	0.82	0.83	0.83	0.83	0.84	0.84	0.85	0.85	0.86	0.86	0.86	0.87	0.87	0.87	0.88	0.88	0.88	0.89
	-0.54	0.45	0.58	0.66	0.72	0.76	0.76	0.77	0.78	0.79	0.80	0.81	0.82	0.82	0.82	0.83	0.83	0.84	0.84	0.85	0.85	0.86	0.86	0.86	0.87	0.87	0.87	0.88	0.88	0.88	0.89
	-0.39	0.45	0.57	0.66	0.71	0.76	0.76	0.77	0.78	0.79	0.80	0.81	0.82	0.82	0.83	0.83	0.83	0.84	0.84	0.85	0.85	0.86	0.86	0.87	0.87	0.87	0.88	0.88	0.88	0.88	0.89
	-0.26	0.44	0.57	0.65	0.71	0.76	0.76	0.77	0.78	0.79	0.80	0.81	0.82	0.82	0.83	0.83	0.83	0.84	0.84	0.85	0.85	0.86	0.86	0.87	0.87	0.87	0.87	0.88	0.88	0.88	0.89
	-0.11	0.44	0.57	0.65	0.71	0.75	0.76	0.77	0.78	0.79	0.80	0.81	0.82	0.82	0.82	0.83	0.83	0.84	0.84	0.85	0.85	0.86	0.86	0.87	0.87	0.87	0.87	0.88	0.88	0.88	0.89
	0.13	0.43	0.56	0.65	0.71	0.75	0.76	0.77	0.78	0.79	0.80	0.81	0.82	0.82	0.83	0.83	0.84	0.84	0.85	0.85	0.85	0.86	0.86	0.87	0.87	0.87	0.88	0.88	0.88	0.89	0.89
	0.41	0.42	0.55	0.64	0.70	0.75	0.76	0.77	0.78	0.79	0.80	0.81	0.82	0.82	0.83	0.83	0.84	0.84	0.85	0.85	0.86	0.86	0.87	0.87	0.87	0.88	0.88	0.88	0.88	0.89	0.89
	0.66	0.41	0.54	0.63	0.70	0.75	0.76	0.77	0.78	0.79	0.80	0.81	0.82	0.82	0.83	0.83	0.84	0.84	0.85	0.85	0.85	0.86	0.86	0.87	0.87	0.88	0.88	0.88	0.88	0.89	0.89
	0.95	0.39	0.53	0.62	0.69	0.74	0.75	0.77	0.78	0.79	0.80	0.82	0.82	0.83	0.83	0.84	0.84	0.85	0.85	0.85	0.86	0.87	0.87	0.87	0.88	0.88	0.88	0.89	0.89	0.89	0.90
	1.89	0.36	0.50	0.61	0.68	0.74	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80

Table 3.8: Optimal Pricing with Daily Price Changes (Four Seats)

		Days Prior to Game																												
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
	Mean price	**0.50**	**0.65**	**0.74**	**0.80**	**0.84**	**0.89**	**0.90**	**0.91**	**0.92**	**0.89**	**0.94**	**0.94**	**0.95**	**0.95**	**0.95**	**0.96**	**0.96**	**0.97**	**0.97**	**0.97**	**0.98**	**0.98**	**0.98**	**0.99**	**0.99**	**0.99**	**1.00**	**1.00**	**1.00**	**1.00**
	Non NB tickets	0.53	0.69	0.77	0.84	0.88	0.89	0.90	0.91	0.92	0.93	0.93	0.94	0.94	0.95	0.95	0.96	0.96	0.97	0.97	0.98	0.97	0.97	0.98	0.98	0.98	0.99	0.99	0.99	1.00	1.00
	-12.83	0.54	0.69	0.77	0.83	0.87	0.87	0.88	0.89	0.90	0.91	0.91	0.92	0.92	0.93	0.93	0.94	0.94	0.95	0.95	0.96	0.96	0.96	0.96	0.97	0.97	0.97	0.98	0.98	0.98	0.99
	-4.60	0.54	0.69	0.78	0.84	0.88	0.89	0.90	0.91	0.92	0.93	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.96	0.96	0.96	0.96	0.97	0.97	0.97	0.97	0.98	0.98	0.98	0.98	0.99
	-3.32	0.54	0.68	0.77	0.82	0.86	0.87	0.88	0.89	0.90	0.90	0.91	0.91	0.92	0.92	0.92	0.93	0.93	0.93	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96	0.97	0.97	0.97	0.97
	-2.50	0.53	0.68	0.76	0.81	0.85	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.91	0.92	0.92	0.93	0.93	0.93	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96	0.96	0.97
	-2.16	0.53	0.67	0.75	0.81	0.85	0.86	0.87	0.87	0.88	0.89	0.90	0.90	0.91	0.91	0.91	0.92	0.92	0.93	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96	0.97
	-1.94	0.53	0.67	0.75	0.80	0.84	0.85	0.86	0.87	0.88	0.89	0.90	0.90	0.90	0.91	0.91	0.92	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96	0.96
	-1.71	0.52	0.66	0.75	0.80	0.84	0.85	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.92	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96	0.96
	-1.51	0.52	0.66	0.74	0.80	0.84	0.85	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.91	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96	0.96
δ_{NB}	-1.33	0.52	0.66	0.74	0.80	0.84	0.85	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.90	0.91	0.91	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96
	-1.15	0.51	0.65	0.74	0.79	0.84	0.84	0.85	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.92	0.92	0.93	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96
	-0.99	0.51	0.65	0.73	0.79	0.83	0.84	0.85	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.92	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96
	-0.86	0.50	0.65	0.73	0.79	0.83	0.84	0.85	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.92	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.95	0.96	0.96
	-0.71	0.50	0.64	0.73	0.79	0.83	0.84	0.85	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.92	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96
	-0.54	0.50	0.64	0.73	0.79	0.83	0.84	0.85	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.92	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.95	0.96	0.96
	-0.39	0.49	0.64	0.73	0.79	0.83	0.84	0.85	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.92	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96
	-0.26	0.49	0.63	0.72	0.78	0.83	0.84	0.85	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.92	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96
	-0.11	0.48	0.63	0.72	0.78	0.83	0.84	0.85	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.92	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96
	0.13	0.48	0.63	0.72	0.78	0.83	0.84	0.85	0.86	0.87	0.88	0.89	0.89	0.90	0.90	0.91	0.91	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.95	0.96	0.96	0.96
	0.41	0.47	0.62	0.71	0.78	0.83	0.84	0.85	0.86	0.87	0.88	0.89	0.90	0.90	0.91	0.91	0.91	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96	0.97
	0.66	0.45	0.61	0.71	0.77	0.82	0.84	0.85	0.86	0.87	0.88	0.89	0.90	0.90	0.90	0.91	0.91	0.92	0.92	0.93	0.93	0.94	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96	0.97
	0.95	0.43	0.59	0.70	0.77	0.82	0.84	0.85	0.86	0.87	0.88	0.90	0.90	0.91	0.91	0.91	0.92	0.92	0.93	0.93	0.94	0.94	0.95	0.95	0.95	0.96	0.96	0.96	0.97	0.97	0.97
	1.89	0.40	0.57	0.68	0.76	0.82	1.80	1.80	1.80	1.80	0.88	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80

Table 3.9: Optimal Pricing with Daily Price Changes (Two Seats)

		Days Prior to Game																													
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
	Mean price	**0.57**	**0.74**	**0.84**	**0.90**	**0.95**	**1.01**	**1.02**	**1.00**	**1.00**	**1.01**	**1.05**	**1.06**	**1.06**	**1.07**	**1.07**	**1.08**	**1.08**	**1.09**	**1.09**	**1.09**	**1.10**	**1.10**	**1.10**	**1.11**	**1.11**	**1.11**	**1.12**	**1.12**	**1.12**	**1.12**
	Non NB tickets	0.61	0.78	0.88	0.94	0.99	1.01	1.02	1.03	1.04	1.05	1.05	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.09	1.09	1.09	1.09	1.09	1.10	1.10	1.10	1.11	1.11	1.11	1.12
	-12.83	0.61	0.78	0.87	0.93	0.98	1.00	1.01	1.02	1.03	1.04	1.04	1.05	1.05	1.06	1.06	1.07	1.07	1.08	1.08	1.08	1.09	1.09	1.09	1.10	1.10	1.10	1.11	1.11	1.11	1.11
	-4.60	0.61	0.78	0.88	0.94	0.98	1.00	1.01	1.02	1.03	1.04	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.08	1.08	1.08	1.09	1.09	1.09	1.09	1.10	1.10	1.10	1.10	1.11	1.11
	-3.32	0.61	0.77	0.86	0.92	0.96	0.99	1.00	1.01	1.02	1.02	1.03	1.04	1.04	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.08	1.08	1.08	1.08	1.09	1.09	1.09	1.10	1.10	1.10
	-2.50	0.60	0.76	0.85	0.91	0.95	0.98	0.99	1.00	1.01	1.01	1.02	1.03	1.03	1.04	1.04	1.04	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.09	1.09	1.09
	-2.16	0.60	0.76	0.85	0.91	0.95	0.98	0.99	1.00	1.00	1.01	1.02	1.03	1.03	1.03	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.07	1.07	1.07	1.07	1.08	1.08	1.08	1.09	1.09
	-1.94	0.59	0.76	0.85	0.90	0.95	0.97	0.98	0.99	1.00	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.09	1.09
	-1.71	0.59	0.75	0.84	0.90	0.95	0.97	0.98	0.99	1.00	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.04	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.08	1.09
	-1.51	0.59	0.75	0.84	0.90	0.95	0.97	0.98	0.99	1.00	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.04	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.08	1.09
ÎŽNB	-1.33	0.58	0.75	0.84	0.90	0.94	0.97	0.98	0.99	1.00	1.01	1.02	1.02	1.03	1.03	1.03	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.07	1.07	1.07	1.07	1.08	1.08	1.08	1.09
	-1.15	0.58	0.74	0.83	0.90	0.94	0.97	0.98	0.99	1.00	1.01	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.09
	-0.99	0.57	0.74	0.83	0.89	0.94	0.97	0.98	0.99	1.00	1.01	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.09
	-0.86	0.57	0.74	0.83	0.89	0.94	0.97	0.98	0.99	1.00	1.00	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.04	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.08
	-0.71	0.57	0.73	0.83	0.89	0.94	0.97	0.98	0.99	1.00	1.00	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.09
	-0.54	0.56	0.73	0.83	0.89	0.94	0.96	0.98	0.99	0.99	1.00	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.04	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.08
	-0.39	0.56	0.73	0.83	0.89	0.94	0.97	0.98	0.99	1.00	1.00	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.09
	-0.26	0.55	0.73	0.82	0.89	0.94	0.96	0.98	0.99	1.00	1.00	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.09
	-0.11	0.55	0.72	0.82	0.89	0.94	0.96	0.97	0.99	0.99	1.00	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.09
	0.13	0.54	0.72	0.82	0.89	0.94	0.97	0.98	0.99	1.00	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.04	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.09	1.09
	0.41	0.53	0.71	0.82	0.89	0.94	0.97	0.98	0.99	1.00	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05	1.05	1.05	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.08	1.09	1.09
	0.66	0.52	0.70	0.81	0.88	0.93	0.96	0.98	0.99	1.00	1.01	1.02	1.02	1.03	1.03	1.04	1.04	1.05	1.05	1.06	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.09	1.09	1.09
	0.95	0.50	0.69	0.81	0.88	0.93	0.97	0.98	0.99	1.00	1.01	1.03	1.03	1.04	1.04	1.04	1.05	1.05	1.06	1.06	1.07	1.07	1.07	1.08	1.08	1.08	1.09	1.09	1.09	1.09	1.10
	1.89	0.46	0.67	0.80	0.88	0.93	1.80	1.80	0.99	1.00	1.01	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80	1.80

Table 3.10: Price Regressions With Competition Effects

	(1) Broker Listings	(2) Broker Listings
Sample	Full	Full
Average Price 0 to 2 Days Before Game	0.47	0.47
Days-to-go coefficients		
2 to 4 days	0.091***	0.101***
	(0.022)	(0.024)
5 to 7 days	0.192***	0.196***
	(0.027)	(0.027)
8 to 10 days	0.297***	0.310***
	(0.027)	(0.026)
11 to 13 days	0.444***	0.419***
	(0.026)	(0.026)
14 to 16 days	0.463***	0.402***
	(0.025)	(0.027)
17 to 19 days	0.352***	0.269***
	(0.027)	(0.029)
20 to 22 days	0.385***	0.267***
	(0.031)	(0.032)
23 to 28 days	0.397***	0.252***
	(0.029)	(0.030)
29 to 34 days	0.426***	0.232***
	(0.028)	(0.031)
35 or more days	0.481***	0.200***
	(0.026)	(0.035)
Minimum Non-Broker Price	0.311***	-0.015
	(0.051)	(0.038)
No Non-Broker Listings	0.244***	0.025
	(0.043)	(0.034)
DTG* Minimum Non-Broker Price	-	0.007***
		(0.001)
DTG*No Non-Broker Listings	-	0.006***
		(0.001)
Game-face value fixed effects	Y	Y
Row, # of seat and team performance controls	Y	Y
Observations	87,401	87,401
R-squared	0.73	0.76

Note: standard errors in parentheses robust to heteroskedasticity and clustered on the game. ***,** and * denote statistical significance at the 1, 5 and 10% levels.

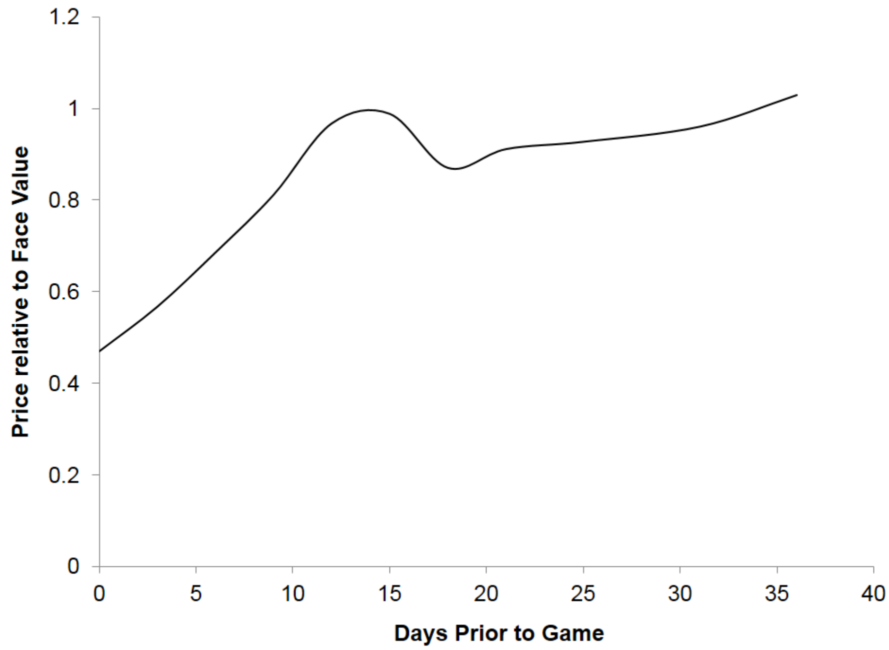


FIGURE 3.1: Broker Prices

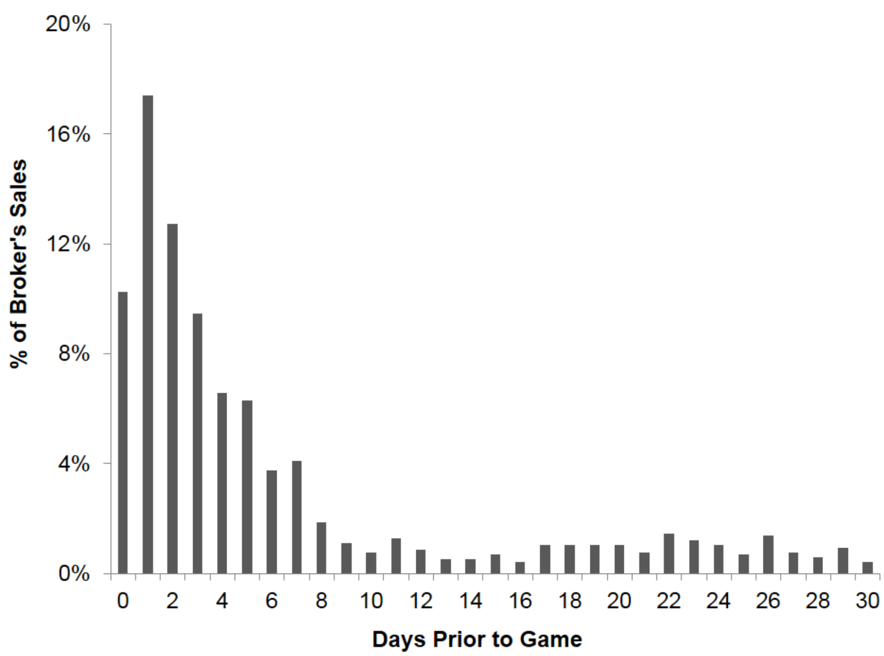


FIGURE 3.2: Timing of Broker Sales

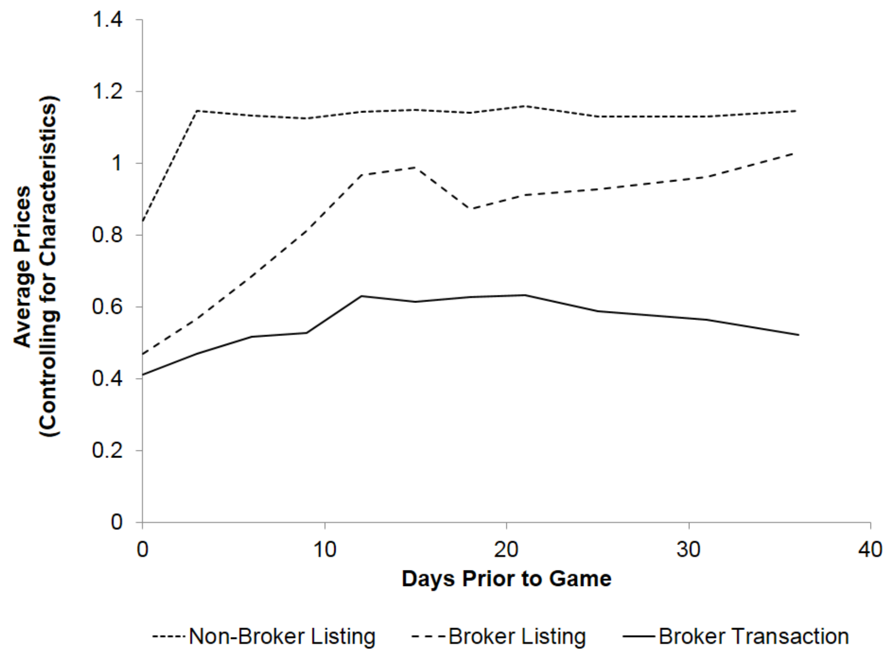


Figure 3.3: Price Paths

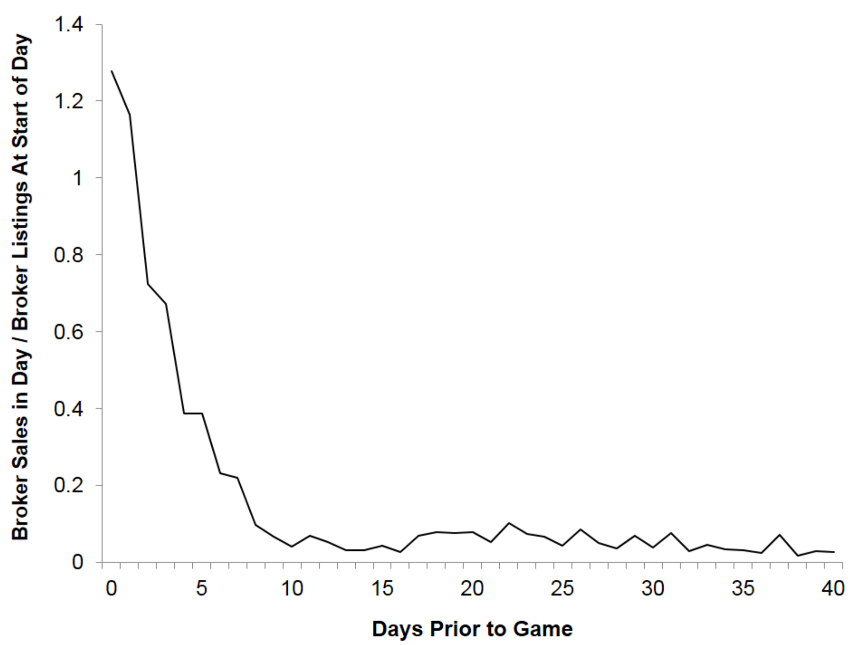


Figure 3.4: Broker Sales Rate

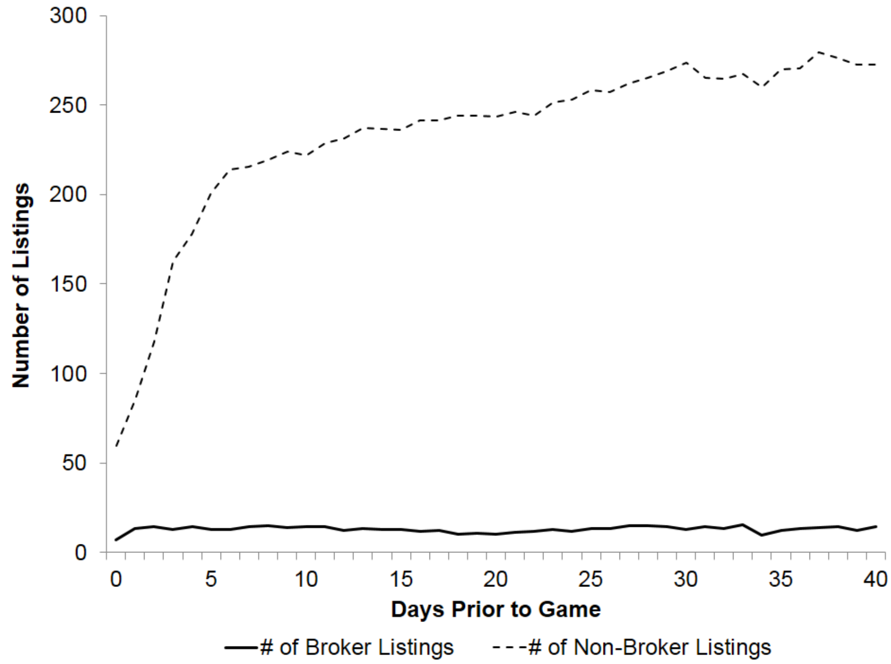
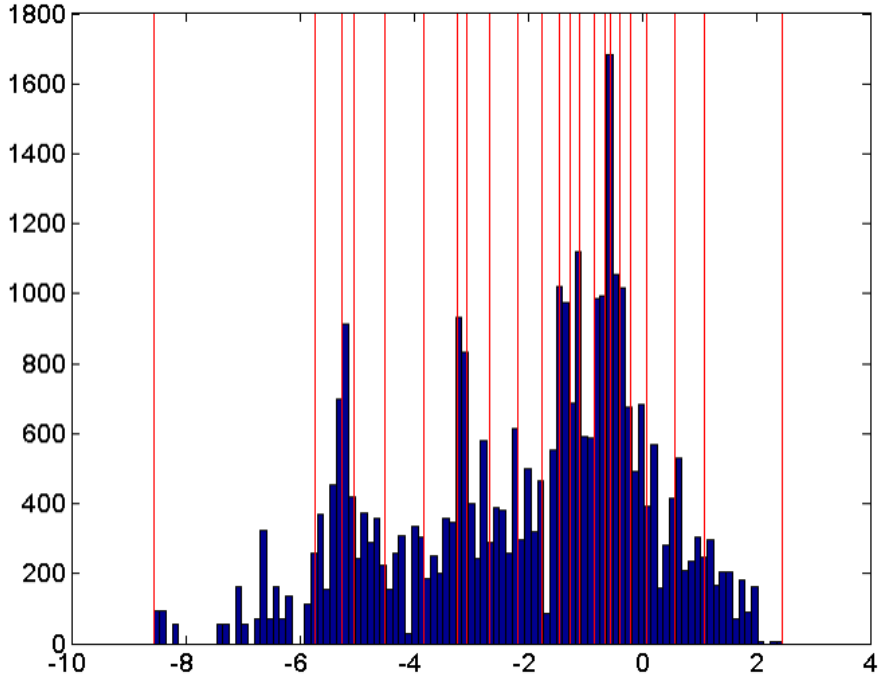
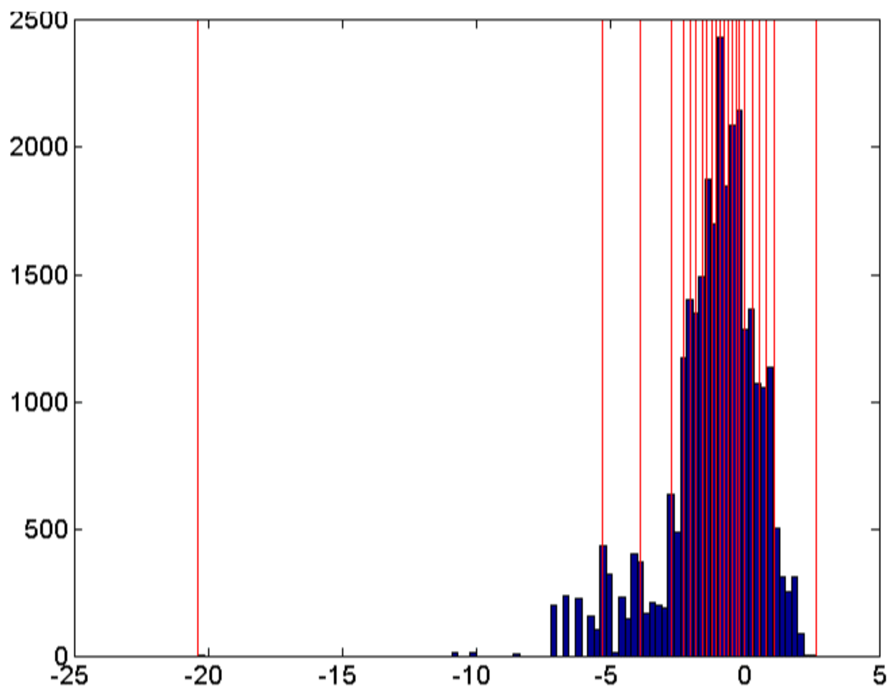


FIGURE 3.5: Average Number of Listings for Game

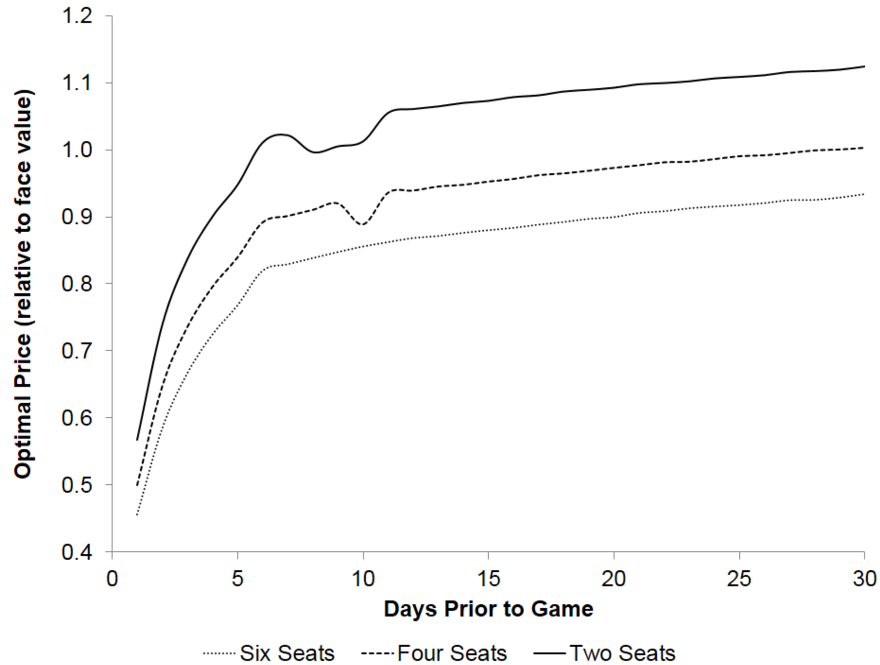


(a) Broker δ

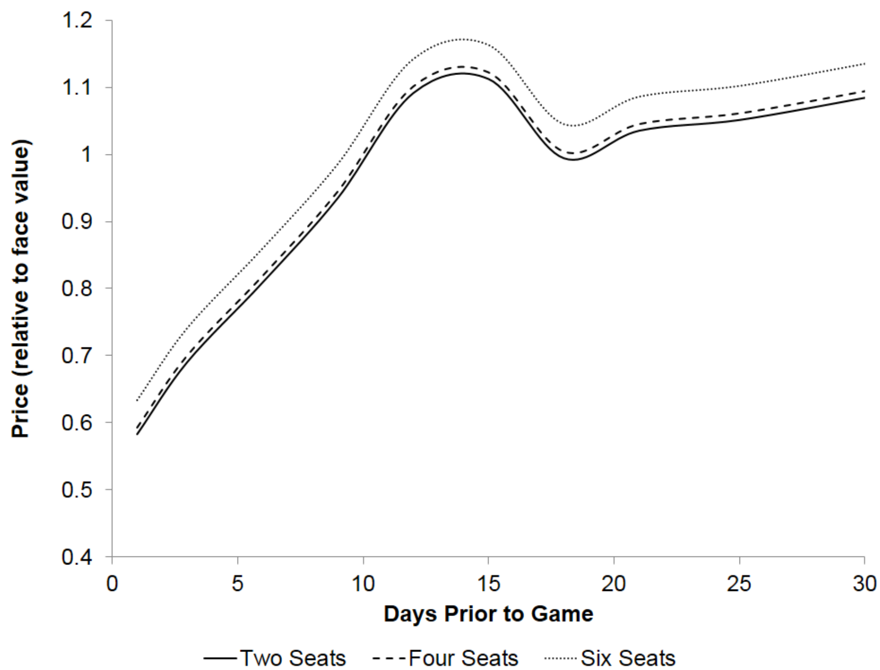


(b) Non-Broker δ

FIGURE 3.6: Bins for δ State Variables



(a) Optimal Prices with Daily Price Changes



(b) Average Price for Row 10 seats

FIGURE 3.7: Optimal Pricing Compared with Actual Average Prices

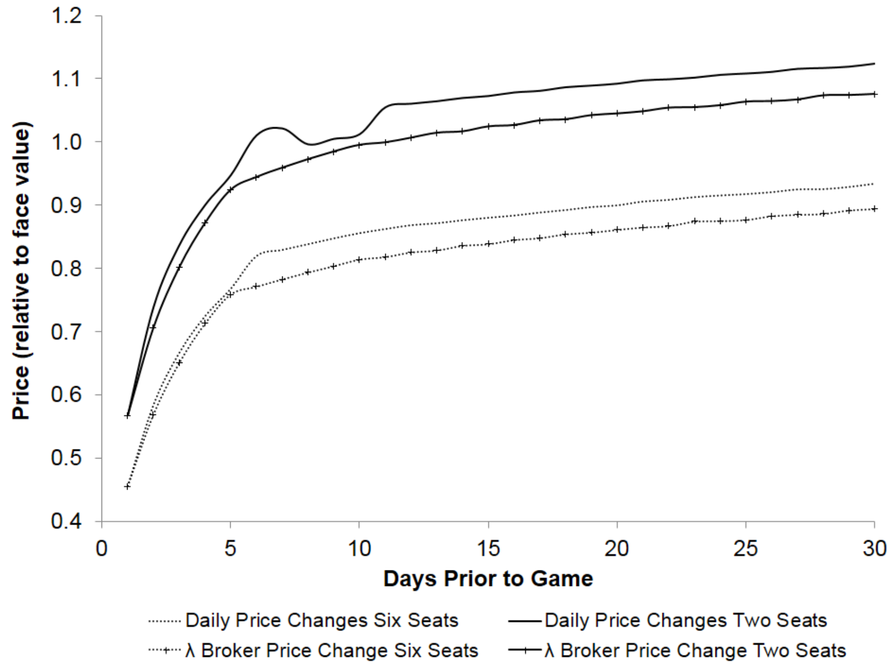


FIGURE 3.8: Optimal Prices with Daily and Estimated Price Setting

Appendix A

Invertibility of Market Share Function in GEV Model

In this appendix, I show the invertibility of market share function discussed in BLP (1995) can be extended to the case of GEV model with a slight modification. They define a function $f : \mathbf{R}^K \to \mathbf{R}^K$. If the following conditions are satisfied, then there is a unique fixed point x_0 to f in $\mathbf{R}^K$.

1. $\forall x \in \mathbf{R}^K$, $f(x)$ is continuously differentiable with $\forall j$ and k, $\partial f_j(x)/\partial x_k \geqslant 0$ and $\sum_{k=1}^{K} \partial f_j(x)/\partial x_k < 1$.
2. $\min_j \inf_x f(x) \equiv x > -\inf$.
3. There is a value, $\bar{x}$, with the property that if for any j, $x_j \geqslant \bar{x}$, then for some k, $f_k(x) < x_k$.

Conditions (2) and (3) are trivial in our case, so I only show the proof for condition (1). Consider $f(\delta) = \delta + \min_l\{\rho_l\} [\log(s) - \log(s(\delta))]$. In our demand model,

$$G(e^{\delta}) = e^{\delta_0} + \sum_l a_l \left[\sum_k \left(\sum_j I(j,k,l) e^{\frac{\delta_j}{\rho_l}} \right)^{\rho_l} \right]$$

and

$$s_j \quad = \quad \frac{1}{G(e^{\delta})} \sum_l a_l e^{\frac{\delta_j}{\rho_l}} \left(\sum_{j'} I(j', k_{jl}, l) e^{\frac{\delta_{j'}}{\rho_l}} \right)^{\rho_l - 1} = \frac{\sum_l a_l e^{\frac{\delta_j}{\rho_l}} T_{jl}^{\rho_l - 1}}{G}.$$

We could show that

$$\frac{\partial f_j}{\partial \delta_j} = 1 + \min\{\rho_l\} s_j - \sum_l \left(\frac{\min\{\rho_l\}}{\rho_l} \right) \left(\frac{a_l e^{\frac{\delta_j}{\rho_l}} T_{jl}^{\rho_l - 1}}{\sum_l a_l e^{\frac{\delta_j}{\rho_l}} T_{jl}^{\rho_l - 1}} \right) \left(1 - (1 - \rho_l) \frac{e^{\frac{\delta_j}{\rho_l}}}{T_{jl}} \right) > 0$$

and for $m \neq j$,

$$\frac{\partial f_j}{\partial \delta_m} = \min\{\rho_l\} \left[s_m - \sum_l \left(\frac{\rho_l - 1}{\rho_l} \right) \left(\frac{a_l e^{\frac{\delta_j}{\rho_l}} T_{jl}^{\rho_l - 1}}{\sum_l a_l e^{\frac{\delta_j}{\rho_l}} T_{jl}^{\rho_l - 1}} \right) \left(\frac{I(m, k_{jl}, l) e^{\frac{\delta_m}{\rho_l}}}{T_{jl}} \right) \right] > 0.$$

In addition,

$$\begin{aligned} \sum_m \frac{\partial f_j}{\partial \delta_m} \quad &= \quad 1 + \min\{\rho_l\} \Bigg[\sum_m s_m - \sum_l \left(\frac{1}{\rho_l} \right) \left(\frac{a_l e^{\frac{\delta_j}{\rho_l}} T_{jl}^{\rho_l - 1}}{\sum_l a_l e^{\frac{\delta_j}{\rho_l}} T_{jl}^{\rho_l - 1}} \right) \\ &\qquad \left(1 - (1 - \rho_l) \frac{\sum_m I(m, k_{jl}, l) e^{\frac{\delta_m}{\rho_l}}}{T_{jl}} \right) \Bigg] \\ &= \quad 1 + \min\{\rho_l\} \left(\sum_m s_m - 1 \right) < 1. \end{aligned}$$

Appendix B

Identifying Broker Listings in the StubHub Data

In this section, we discuss how we identified broker's listings in the StubHub data. From the broker we have a record of sales transactions. This record indicates the number of seats sold, transaction time, the game, section, row, and number of seats, and the platform on which the tickets were sold (usually StubHub). From StubHub we have a record of all the tickets listed (by anyone, not just the broker) for each game-section-row at various points in time. The data was gathered from the StubHub website approximately every 3 hours from December 2009 to April 2010.

Since the listings on StubHub are anonymous, we need to identify broker's listings in the StubHub data based on the available information. To accomplish this we first generated a variable called *BrokerTrans*, which is equal to one for a listing when tickets disappear from that listing at the next download and the timing and number of seats match with a sale recorded in the broker data. Then, we use logical tests in Stata to create categories for the StubHub listings. Below is a description for each category. Each description includes an explanation of the logical test performed and the rationale behind the category.

1. *BrokerTicket1*: This variable identifies StubHub listings which are the only ones in a game-section-download that match the number of seats held by the broker at the time of the listing. The variable equals one if the number of seats match, and zero otherwise. By definition, the listings with *BrokerTrans* equal to one must have *BrokerTicket1* equal to one as well.

2. *BrokerTicket2*: This variable works backward by recognizing when a broker's transaction has occurred, and then identifying the listings leading up to the transaction. If we find a listing for a game-section-row the only one that matches the broker data and *BrokerTicket1* equals zero, then *BrokerTicket2* equals one for this listing and for listings with the same number of seats when they appear at earlier download times. *BrokerTicket2* is less restrictive than *BrokerTicket1* in that it does not require the number of seats on the listing to match the number held by the broker. This accounts for the possibility that the broker did not list all tickets together simultaneously on StubHub.

3. *ProbableBrokerTicket1*: This variable follows *BrokerTicket1* except it identifies multiple listings from the same download time with the same number of seats. For instance, if the broker has four seats for a particular game-section-row and there are two listings of four seats on StubHub, *BrokerTicket1* equals zero for both listings since it cannot distinguish between the two identical listings. In this case *ProbableBrokerTicket1* equals one for both listings.

4. *ProbableBrokerTicket2*: This variable follows *BrokerTicket1*, but it allows for the possibility that the order of transactions from the broker data is different from the order on StubHub.

5. *ProbableBrokerTicket3*: This variable is equal to one whenever the number of

tickets held by the broker for a particular game-section-row is greater than the total number of tickets listed for that game-section-row on StubHub and the listing has *BrokerTicket2* equal to zero.

6. *PossibleBrokerTicket1*: This variable identifies tickets that are continuously listed except for a single gap of 1 to 4 downloads (i.e. for a period of time the tickets disappear but then reappear with the same price and the same number of seats). This test is only applied to listings where *BrokerTicket1* or *BrokerTicket2* equals one later.

7. *PossibleBrokerTicket2*: This variable is similar to *PossibleBrokerTicket1* except the listing price must vary across the gap (number of seats must remain the same).

8. *PossibleBrokerTicket3*: This variable is the same as *PossibleBrokerTicket1* except the gap length is longer (5 or more downloads).

9. *PossibleBrokerTicket4*: The same as *PossibleBrokerTicket2* except the gap length is longer (5 or more downloads).

Note that due to technical reasons, some broker transactions were not identified in the StubHub data. For example, a broker transaction which happened when our download system was down would not be found in the StuhHub data. Also, the broker recorded transactions taking place on holidays in the morning of the business days right after the holidays. This made the transaction time from the broker not consistent with the time when the listings disappeared from StubHub during the holidays. Table B.1 summarizes the results of the matching process.

Table B.1: Matching Datasets

Category	Number of Observations	Percentage
BrokerTicket1	87,401	5.67%
BrokerTicket2	15,901	1.03%
ProbableBrokerTicket1	9,118	0.59%
ProbableBrokerTicket2	0	0.00%
ProbableBrokerTicket3	29,968	1.94%
PossibleBrokerTicket1	378	0.02%
PossibleBrokerTicket2	374	0.02%
PossibleBrokerTicket3	4	0.00%
PossibleBrokerTicket4	9	0.00%
Non-Broker	1,399,446	90.72%
Total	1,542,599	100.00%

Bibliography

Allenby, G. M. and Lenk, P. J. (1995), "Reassessing brand loyalty, price sensitivity, and merchandising effects on consumer brand choice," *Journal of Business & Economic Statistics*, 13, 281–289.

Arcidiacono, P., Bayer, P., Blevins, J. R., and Ellickson, P. B. (2012), "Estimation of dynamic discrete choice models in continuous time with an application to retail competition," Tech. rep., National Bureau of Economic Research.

Arcidiacono, P., Ellickson, P. B., Landry, P., and Ridley, D. B. (2013), "Pharmaceutical followers," *International Journal of Industrial Organization*, 31, 538–553.

Bajari, P., Benkard, C. L., and Levin, J. (2007), "Estimating dynamic models of imperfect competition," *Econometrica*, 75, 1331–1370.

Benkard, C. L., Bodoh-Creed, A., and Lazarev, J. (2010), "Simulating the dynamic effects of horizontal mergers: Us airlines," *Manuscript, Yale University.*

Berry, S., Levinsohn, J., and Pakes, A. (1995), "Automobile prices in market equilibrium," *Econometrica: Journal of the Econometric Society*, pp. 841–890.

Berry, S., Levinsohn, J., and Pakes, A. (2004), "Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Car Market," *Journal of Political Economy*, 112, pp. 68–105.

Berry, S., Carnall, M., and Spiller, P. T. (2006), "Airline hubs: Costs, markups and the implications of customer heterogeneity," *Competition policy and antitrust.*

Bitran, G. R. and Mondschein, S. V. (1997), "Periodic pricing of seasonal products in retailing," *Management Science*, 43, 64–79.

Bresnahan, T. F. (1987), "Competition and collusion in the American automobile industry: The 1955 price war," *The Journal of Industrial Economics*, pp. 457–482.

Bresnahan, T. F., Stern, S., and Trajtenberg, M. (1997), "Market Segmentation and the Sources of Rents from Innovation: Personal Computers in the Late 1980s," *The RAND Journal of Economics*, 28, pp. S17–S44.

Bronnenberg, B. J., Dubé, J.-P. H., and Gentzkow, M. (2012), "The Evolution of Brand Preferences: Evidence from Consumer Migration," *American Economic Review*, 102, 2472–2508.

Caro, F. and Gallien, J. (2012), "Clearance pricing optimization for a fast-fashion retailer," *Operations Research*, 60, 1404–1422.

Corts, K. S. (1998), "Third-degree price discrimination in oligopoly: all-out competition and strategic commitment," *The RAND Journal of Economics*, pp. 306–323.

Coscelli, A. (2000), "The importance of doctors' and patients' preferences in the prescription decision," *The Journal of Industrial Economics*, 48, 349–369.

Crawford, G. S. and Shum, M. (2005), "Uncertainty and learning in pharmaceutical demand," *Econometrica*, 73, 1137–1173.

Dalen, D. M., Furu, K., Locatelli, M., and Strøm, S. (2011), "Generic substitution: micro evidence from register data in Norway," *The European Journal of Health Economics*, 12, 49–59.

Dubé, J.-P., Hitsch, G. J., and Rossi, P. E. (2010), "State dependence and alternative explanations for consumer inertia," *The RAND Journal of Economics*, 41, 417–445.

Ellison, G. and Ellison, S. F. (2007), "Strategic entry deterrence and the behavior of pharmaceutical incumbents prior to patent expiration," Tech. rep., National Bureau of Economic Research.

Ellison, G. and Ellison, S. F. (2009), "Search, obfuscation, and price elasticities on the internet," *Econometrica*, 77, 427–452.

Elmaghraby, W. and Keskinocak, P. (2003), "Dynamic pricing in the presence of inventory considerations: Research overview, current practices, and future directions," *Management Science*, 49, 1287–1309.

Erdem, T. (1996), "A dynamic analysis of market structure based on panel data," *Marketing Science*, 15, 359–378.

Gallego, G. and Hu, M. (2014), "Dynamic pricing of perishable assets under competition," *Management Science*.

Gallego, G. and Van Ryzin, G. (1994), "Optimal dynamic pricing of inventories with stochastic demand over finite horizons," *Management science*, 40, 999–1020.

Heching, A., Gallego, G., and van Ryzin, G. (2002), "Mark-down pricing: An empirical analysis of policies and revenue potential at one apparel retailer," *Journal of Revenue and Pricing Management*, 1, 139–160.

Holmes, T. J. (1989), "The effects of third-degree price discrimination in oligopoly," *American Economic Review*, 79, 244–250.

Jayawardhana, J. (2013), "Direct-to-consumer advertising and consumer welfare," *International Journal of Industrial Organization*, 31, 164–180.

Keane, M. P. (1997), "Modeling heterogeneity and state dependence in consumer choice behavior," *Journal of Business & Economic Statistics*, 15, 310–327.

Kincaid, W. M. and Darling, D. A. (1963), "An inventory pricing problem," *Journal of Mathematical Analysis and Applications*, 7, 183–208.

Klemperer, P. (1995), "Competition when consumers have switching costs: An overview with applications to industrial organization, macroeconomics, and international trade," *The Review of Economic Studies*, 62, 515–539.

Krishnamurthi, L. and Raj, S. P. (1991), "An empirical analysis of the relationship between brand loyalty and consumer price elasticity," *Marketing Science*, 10, 172–183.

Li, J., Granados, N., and Netessine, S. (2011), "Are consumers strategic? Structural estimation from the air-travel industry," Tech. rep., Working Paper.

Lin, K. Y. and Sibdari, S. Y. (2009), "Dynamic price competition with discrete customer choices," *European Journal of Operational Research*, 197, 969–980.

Lundin, D. (2000), "Moral hazard in physician prescription behavior," *Journal of Health Economics*, 19, 639–662.

McAfee, R. P. and Te Velde, V. (2006), "Dynamic pricing in the airline industry," *forthcoming in Handbook on Economics and Information Systems, Ed: TJ Hendershott, Elsevier.*

McAfee, R. P. and Velde, V. (2008), "Dynamic pricing with constant demand elasticity," *Production and Operations Management*, 17, 432–438.

McFadden, D. (1973), "Conditional logit analysis of qualitative choice behavior," .

McFadden, D. et al. (1978), *Modelling the choice of residential location*, Institute of Transportation Studies, University of California.

Narasimhan, C. (1984), "A price discrimination theory of coupons," *Marketing Science*, 3, 128–147.

Nevo, A. (2000), "A Practitioner's Guide to Estimation of Random-Coefficients Logit Models of Demand," *Journal of Economics & Management Strategy*, 9, 513–548.

Nevo, A. and Rossi, F. (2008), "An approach for extending dynamic models to settings with multi-product firms," *Economics Letters*, 100, 49–52.

Nevo, A. and Wolfram, C. (2002), "Why do manufacturers issue coupons? An empirical analysis of breakfast cereals," *RAND Journal of Economics*, pp. 319–339.

Osborne, M. (2011), "Consumer learning, switching costs, and heterogeneity: A structural examination," *Quantitative Marketing and Economics*, 9, 25–70.

Puller, S. L., Sengupta, A., and Wiggins, S. N. (2009), "Testing Theories of Scarcity Pricing in the Airline Industry," Tech. rep., National Bureau of Economic Research.

Ryan, S. P. (2012), "The costs of environmental regulation in a concentrated industry," *Econometrica*, 80, 1019–1061.

Smith, B. C., Leimkuhler, J. F., and Darrow, R. M. (1992), "Yield management at American airlines," *Interfaces*, 22, 8–31.

Soysal, G. P. and Krishnamurthi, L. (2012), "Demand dynamics in the seasonal goods industry: An empirical analysis," *Marketing Science*, 31, 293–316.

Sweeting, A. (2012), "Dynamic pricing behavior in perishable goods markets: Evidence from secondary markets for major league baseball tickets," *Journal of Political Economy*, 120, 1133–1172.

Thompson, D. F. (1993), "Understanding financial conflicts of interest," *New England Journal of Medicine*, 329, 573–573.

Vulcano, G., van Ryzin, G., and Chaar, W. (2010), "OM Practice-Choice-Based Revenue Management: An Empirical Study of Estimation and Optimization," *Manufacturing & Service Operations Management*, 12, 371–392.

Wooldridge, J. M. (2010), *Econometric analysis of cross section and panel data*, MIT press.

Xu, X. and Hopp, W. J. (2006), "A monopolistic and oligopolistic stochastic flow revenue management model," *Operations Research*, 54, 1098–1109.

Zhao, W. and Zheng, Y.-S. (2000), "Optimal dynamic pricing for perishable assets with nonhomogeneous demand," *Management science*, 46, 375–388.

Biography

Chung-Ying Lee was born on February 12, 1982 in Taipei, Taiwan. He attended National Taiwan University from 2000 to 2004, where he received a B.B.A. in Finance and was ranked second out of 138 in his graduating class. In his senior year, he went to the University of Pennsylvania, Philadelphia, Pennsylvania as an exchange student and earned Dean's List 2003-2004. After completing the military service in Taiwan, he started to study in the Economics M.A. program at Duke University, Durham, North Carolina in 2007. He went on to study in the doctoral program at Duke University and earned a Ph.D. in Economics in 2014. Chung-Ying will be starting as an Assistant Professor in Economics at National Taiwan University in the Fall of 2014.